DOCUMENT RESUME

ED 455 271 TM 033 074

AUTHOR Zhang, Liru
TITLE Delaware Student Testing Program: Report on Special Writing Study.
INSTITUTION Delaware State Dept. of Education, Dover. Assessment and Accountability Branch.
PUB DATE 2000-10-00
NOTE 113p.; For a report on the comparison of Delaware Student Testing Program results for 1998 through 2000, see TM 033 073.
PUB TYPE Numerical/Quantitative Data (110) -- Reports - Research (143)
EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Academic Achievement; Elementary Education; Elementary School Students; *Low Achievement; *Scores; Scoring; State Programs; Teachers; *Test Results; *Testing Programs; Writing Achievement; *Writing Tests
IDENTIFIERS *Delaware Student Testing Program

ABSTRACT

This study invesitigated possible reasons for the low performance on the text-based writing assessment of the Delaware Student Testing Program (DSTP) in 2000, especially for grades 3 and 5, and considered ways to improve classroom instruction. In the first part of the study, a panel of teachers reviewed the anchor papers from the assessment and the process of testing. Panel members rescored the anchor papers of a given grade independently and then worked in a small group to discuss and finalize their scores. In the second part of the study, a second panel of teachers participated in a rescoring session for a sample of 100 text-based writings per grade, scoring students' writings holistically and analytically using the five-point scoring rubric. Both panels also reviewed the directions for administering the test and test booklets. Results of the study are reported in the categories of: (1) the process of testing; (2) construct validity evidence; (3) text-based writing scoring; (4) text-based writing development; and (5) text-based writing instruction. In each of these categories, some possible explanations for student performance were identified, including the introduction of a new text-based writing task at each grade level in 2000, as exemplified by the change from persuasive writing to informative writing at grade 3. The paper cautions that the study was done rapidly and that much additional research is needed, focusing on the stability of scores over time and the processes of reader training and scoring. Attachments contain the scoring rubric, instructions for the anchor paper review, rescoring directions, a list of panel members, correlation matrices, and records for the rescoring. (Contains 14 tables, 4 figures, and 8 references.) (SLD)

TM

Delaware Student Testing Program

Report on Special Writing Study

Prepared by the Assessment and Analysis Group
Assessment and Accountability Branch
Delaware Department of Education

October 2000

Officers of the Delaware Department of Education

State Board of Education

James L. Spartz, Ed. D., President
Jean W. Allen, Vice President
Mary B. Graham, Esquire
John W. Jardine, Jr.
Joseph A. Pika, Ph. D.
Dennis J. Savage
Claiboune D. Smith, Ph. D.

Valerie A. Woodruff
Secretary of Education
Jennifer W. Davis
Deputy Secretary of Education
David Blowman
Executive Assistant

Robin R. Taylor, M. Ed
Acting Associate Secretary, Assessment and Accountability Branch
Nancy J. Wilson, Ph. D.
Associate Secretary, Curriculum and Instructional Improvement Branch

Wendy B. Roberts, Ph. D.
Acting Director, Assessment and Analysis Group
Darlene J. Bolig, Ed. D.
Jeffery Fleming, M. S.
James F. Hertzog, M. Ed.
Nancy Maihoff, Ph. D.
Jon Manon. Ph. D., University of Delaware
Joann F. Prewitt, M. A.
Julie A. Schmidt, Ph. D., University of Delaware
Carole D. White, M. B. A.
Liru Zhang, Ph. D.

Qi Tao, Ph. D.
Technology Management and Design
Policy and Administrative Service Branch

Support Staff:

Krista D. Holloway	Erin L. Pieshala
Barbara F. O'Neal	Kimberly K. Rodriguez

The report was prepared by Liru Zhang

The panel comments on Test Development in the section of Results of the Study was summarized by Darlene Bolig.

The panel comments on Classroom Instruction in the section of Results of the Study was summarized by Mike Kelly.

This report was reviewed and edited by Robin Taylor, Wendy Roberts, Darlene Bolig, and Mike Kelly.

Report on a Special Writing Study

Introduction

The objective of the Delaware Student Testing Program (DSTP) is to measure student progress toward the Delaware Content Standards. Each spring, all public school students in grades 3, 5, 8, and 10 take the statewide assessment in reading, writing, and mathematics. The writing assessment consists of a text-based writing task and a stand-alone writing prompt. The text-based writing task links to a passage in the DSTP reading assessment and students' responses to this task are scored twice, once for a reading score and once for a writing score. Both stand-alone and text-based writings are untimed. Students usually take approximately 2 hours, including a 30 minute pre-writing session, to develop, organize, draft, and finalize their stand-alone writings. Only the final draft of this prompt is scored.

A 5-point scoring rubric (Please see Attachment A) is used to score both the text-based and stand-alone responses. One reader scores the text-based writing; two readers score the stand-alone writing. The lowest score for the text-based writing is 1 and the highest possible score is 5; the lowest score for the stand-alone writing is 2 and the highest possible score is 10. The total writing raw score is the sum of the text-based writing score and the stand-alone writing score with the lowest score of 3 and the highest possible score of 15.

Over the past three years, the overall writing scores have declined in grades 3 and 5, remained steady in grade 8, and increased slightly in grade 10 (See Tables 1a -1d). The average performance on the stand-alone writing shows a consistent pattern of increase across years for students in grades 8 and 10; minor fluctuations over time for students in grades 3 and 5. Student performance on the text-based writing, however, dropped to the lowest level in 2000 for all grades except grade 10, where the average scores dropped by .66 from 1999 and .63 from 1998 to 2000 in grade 3; dropped by .59 from 1999 and .76 from 1998 to 2000 in grade 5; and dropped by .28 from 1999 and .45 from 1998 to 2000 in grade 8. Because of the drop in text-based writing scores, the Assessment and Analysis Group decided to conduct a special writing study to investigate the possible reasons for the low performance in 2000, especially in grades 3 and 5.

Purpose of the Study

The primary purposes of this study were (1) to investigate the possible reasons for the low performance on the text-based writing in 2000, especially in grades 3 and 5 and (2) to investigate ways to improve classroom instruction in writing.

Methods of the Study

General Design Due to the time constraints and the availability of information/data, this study focused on the following five aspects:

- Review the test process (i.e., review of test administration and testing materials);
- Review text-based writing scores (i.e., review anchor papers and re-score a sample of students' responses to the text-based writing task, and compare the statistics of text-based writing scores from the field test with and the 2000 DSTP);
- Examine construct validity evidence (i.e., review available data and conduct additional statistical analyses);
- Make recommendations for the development of text-based writing tasks; and
- Make recommendations on ways to improve classroom instruction in writing.

This study included two parts. In Part One, a panel of teachers reviewed the anchor papers and the process of testing (See Attachment B). Anchor papers are a sample of students' writings that are used as benchmarks in scoring. Each anchor paper represents a score point and usually represents the upper and lower levels within each score point. In this study, the panel members re-scored the anchor papers of a given grade independently, and then worked in a small group to discuss and finalize their scores. In Part Two, a second panel of teachers participated in a re-scoring session for a sample of 100 text-based writings per grade. They scored students' writings holistically and analytically using the 5-point scoring rubric. Each writing sample was evaluated by up to 5 teachers. Then, the panel members discussed related issues in test administration, test development, scoring, and classroom instruction (See Attachment C).

Sample of Student Writings A random sample of 100 student responses to the text-based writing task was selected from the population of each grade for re-scoring in this study.

Panels of Teachers Two panels of teachers were invited to participate in this study, one for anchor paper review and one for the re-scoring session (See Attachment D). These teachers were selected based on their expertise in writing, teaching experience, experience in the development of writing assessment and scoring, familiarity with the Delaware Content Standards in English language arts and the writing scoring rubric, geographic location, and availability.

The Anchor Paper Review Panel consisted of 9 members, including 7 teachers (78%) from 7 school districts and 2 staff members from the Department of Education. Among them, 7 are females and 2 males. Seven of the panel members (78%) have served on the test development committees and 2 (22%) were involved in the anchor paper pulling for the 2000 DSTP writing assessment.

The Re-Scoring Panel included 22 members, 20 of them were teachers (91%) from 12 school districts, 1 from the University of Delaware, and 2 from the Department of Education. Forty-one percent of the members (9) have served on the test development committees and nearly half (about a quarter) were involved in the anchor paper pulling for the 2000 DSTP writing assessment.

Data Analysis and Summary of Comments To investigate the possible reasons for the low performance on the text-based writing in 2000, teachers reviewed, discussed, and made recommendations for improving test administration, test development, scoring, and text-

based writing instruction. All comments and subjective evaluations are summarized in attachments. These comments are reported in the section of the "Results of the Study". The results of data analyses are presented in tables and charts. Data analyses for this study include:

- Three-year comparisons of students' writing scores by grade
- Three-year comparisons of statistics of writing scores by gender and grade
- Correlation analysis of all types of writing scores and reading scores by grade for 2000 DSTP
- Comparisons of means and frequency distributions of text-based scores between the 2000 DSTP and field test by grade
- Comparisons of means and frequency distributions of text-based scores from the field test, the 2000 DSTP, and re-scoring
- The discrepancy of text-based scores from the 2000 DSTP and re-scoring
- The reliability indicators of re-scoring the text-based writing

Results of the Study

The results of the study are reported in five categories, the process of testing, construct validity evidence, text-based writing scoring, text-based writing development, and text-based writing instruction.

Process of Testing Both panels reviewed the 2000 DSTP Directions for Administering the Test and test booklets and compared those directions with the previous years' testing materials and the process of testing. Comments on the process of testing focus on the following issues:

1. Two text-based writing prompts, one for field test and one for operational test, should not be given on the same day, especially for younger students.
2. The text-based writing should be given in the beginning of the reading test rather than as the last item of the day.
3. The instructions for the text-based writing should be written to draw students' attention, such as bolded for emphasis, using separate pages to ensure that students understand this item will be scored twice for both reading and writing.
4. The text-based writing task should be formatted similar to the stand-alone writing prompt, such as using pre-writing.

Review of Text-based Writing Scores

Construct Validity Evidence The statistics of the three writing scores, text-based, stand-alone, and the writing total raw scores are compared by gender and grade for 1998, 1999, and 2000 (Table 2). The data show that female students consistently outperformed male students on both text-based and stand-alone writings across all four grades and three years. A slight, but consistent decrease of gender differences in text-based scores for students in grade 3 is observed (Diff score=.30 in 1998; .22 in 1999; .13 in 2000) from the current analysis.

As indicated earlier in this report, the text-based writing tasks attach to a passage in the DSTP reading assessment. This passage includes several multiple-choice (MC) and constructed-response (CR) items. The student's response to one of the CR items was scored as part of the reading score using the reading scoring rubric and the text-based writing score using the writing rubric.

The multiple-choice item has a stem asking a question with four choices, from which students select the best answer to the question. Multiple-choice items are scored dichotomously. Constructed-response items include short answer and extended constructed-response items. Short answer items allow students to make decisions by constructing brief responses that demonstrate the students' understanding of the text. Short answer items are scores on a 0-2 scale. Extended constructed-response items allow students to make decisions by constructing more lengthy responses that demonstrate the students' understanding of the text and require students to provide justification for their responses. The extended constructed-response items are scores on a 0-4 scale. The item used as the text-based writing task is an extended constructed-response item.

Tables 3a and 3b present the correlation coefficients among five reading scores, three writing scores, and the SAT9 reading comprehension test scores by grade from the 2000 DSTP. The analyses are based on the following eight variables:

- MCITEM: The multiple-choice item score is the sum of scores on all MC items attached to the reading passage
- CRITEM: The constructed-response item score is the sum of scores on all constructed-response items attached to the reading passage
- PASSAGE: The passage score is the sum of scores on all MC and CR items attached to the reading passage
- IREADING: The reading item score is the score on the extended constructed-response item attached to the reading passage that was used as the text-based writing task
- TEXT: The text-based writing score is the score on the extended constructed-response item attached to the reading passage that was used as the text-based writing task
- PROMPT: The writing score on the stand-alone writing prompt
- WRITING: The total writing raw score that is the sum of the text-based and stand-alone writing scores
- READING: The DSTP reading score
- SAT9: The reading score on the 30-item SAT9 reading comprehension test

The results show that the correlation coefficients between the text-based writing scores (TEXT) and the item reading scores (IREADING) from the same CR items are .22 for grade 3, .45 for grade 5, .57 for grade 8, and .60 for grade 10. First, the statistics indicate a grade pattern, from the lowest value of the correlation coefficient in grade 3 to the highest value in grade 10. Second, the low correlation in grade 3 suggests that only 5% of the variance from one score associates with the other score; in grade 5, about 20% of

the variance from one score associates with the other score. The correlations between the text-based writing scores (TEXT) and the scores on the MC items (MCITEM) and the CR items (CRITEM) from the reading passage, and the passage scores (PASSAGE) are .19, .31 and .31 in grade 3, which is the lowest among the four grades. Again, the low correlation in grade 3 suggests that only 4% to 10% of the variance of text-based writing scores is associated with the MC item scores, CR item scores, and the passage scores, respectively. Similarly, the correlations between the text-based writing score and the scores on MC items (MCITEM) and CR items (CRITEM) from the reading passage, and the score of the reading passage (PASSAGE) are .26 to .44 in grade 5, which indicate that 7% to 19% of the variance from the text-based writing scores can be accounted for by the scores from reading. Statistics from the current analysis suggest that the text-based writing score is somewhat independent of the reading passage, especially in grades 3 and 5.

The correlations between the text-based writing (TEXT) and stand-alone writing score (PROMPT) range from .36, .41, .41, and .48 for grades 3, 5, 8, and 10, respectively. A grade pattern is observed, where the correlation coefficient for grade 3 is the lowest among the four grades. The low and moderately low correlations across grades suggest that about 13% to 23% of the variance from the text-based writing scores is associated with the stand-alone writing scores. Statistics appear to suggest that the text-based writing measures different types of writing skills or different constructs from the stand-alone writing.

The correlations between SAT9 reading scores, a standardized test, and the DSTP reading scores (READING) are stable across grades, ranging from .84 to .86, and no grade pattern is found. Moreover, the sizes of the correlation coefficients between the SAT9 reading and the stand-alone writing scores (PROMPT) are very close, ranging from .41 to 48 across grades without a grade pattern. The correlations between the SAT9 reading and the text-based writing scores (TEXT), however, show a grade pattern with the lowest coefficient in grade 3 (r=.33) and the highest coefficient in grades 8 and 10 (r=.48). The statistics indicate that there may be more measurement errors involving in the text-based writing scores than that in the stand-alone writing scores, especially in the lower grades.

The correlation matrix among different types of writing scores and reading scores for the 1998 and 1999 DSTP provides additional information for the construct validity (See Attachment D). The correlation coefficients between reading and writing scores are consistent in 1998 and 1999. The correlations between text-based writing and reading scores are higher in 1998 and 1999 (r= .56 in 1998 and r=.60 in 1999 for grade 3; r=.60 in 1998 and r=.56 in 1999 for grade 5) than that in 2000 (r=.33 in grade 3; r=.44 in grade 5). In grade 3, the correlation between text-based and stand-alone writings is lower in 2000 (r=.36) than the previous years (r=.45 in 1998; r=.46 in 1999). The correlation between text-based writing and reading scores also shows the lowest value in 2000 for grade 3 (r=.63 in 1998; r=.68 in 1999; r=.53 in 2000). Such variations of the statistics across years of testing may be due to one or more of the following reasons:

- Low generalization of writing scores across topics, the purposes of writing tasks, and occasions;
- More errors in text-based writing than standard-alone writing in 2000 than the previous years;
- More errors in scoring text-based writing than in scoring stand-alone writing because of using one reader;
- Variations in the characteristics of the reading passages and attached items from year to year and from grade to grade; or
- Variations in writing skills among student populations from year to year.

Review of Text-Based Writing Scores

Re-scoring Anchor Papers To examine the accuracy of scoring, the anchor papers were reviewed and re-scored by the first panel. Anchor papers are typical writing samples that represent each score point and used as benchmarks in reader training and scoring. Usually, writing samples are selected to represent the upper and lower levels of each score point to facilitate the process of scoring. The panel members reviewed, ranked, and re-scored each anchor paper independently. Then, they worked in small groups to discuss their scores in order to achieve an agreement, if possible. The mean of the 4 - 5 new scores from re-scoring process were calculated and compared with the original scores. If the different score between the new score and the original score (Diff = re-score - original score) was equal or greater than a half point (Diff = >.5), this paper was flagged as inconsistency in scoring. The results of re-scoring anchor papers show (See Attachment F) that the new scores and the original scores are highly consistent in grades 3 and 5 (92%), and moderately high in grade 8 (86%) and grade 10 (73%). According to the panel recommendation, three papers should not be used as anchor papers in the future.

Comparing Test Statistics The mean and the relative frequency distributions of text-based writing scores from the field test and the 2000 DSTP were compared. The discrepancies of scores summarized in Table 4 show that the average scores from the field test are consistently higher than the average scores from the 2000 DSTP across the four grades, where the different scores (2000 DSTP score - field test score) are -1.03, -.68, -.26, and -.65 for grades 3, 5, 8, and 10, respectively. Chart 1 illustrates the discrepancy between the relative frequency distributions of the text-based scores from the two occasions. In grade 8, the relative frequency distributions of the writing scores from the field test and the DSTP 2000 are nearly identical with the same mode above the 2-point on the score scale. The difference in mean scores is due to 15% more students receiving a 2-point and 11% less students receiving a 3-point in the 2000 DSTP than that in the field test. Similarly in grade 10, because of the mode of 3-point from the field test and 2-point from the 2000 DSTP, the frequency distribution of the 2000 DSTP shift to left from the frequency distribution of the field test with a lower average score. In grade 3, the frequency distribution of the field test scores approaches a normal distribution with the mode of 3-point; while the frequency distribution of the 2000 DSTP scores is positively skewed with the mode of 1-point, which is 2 score points lower than the mode of the score distribution from the field test. In grade 5, the mode of the writing scores is 3-point of the field test scores with a nearly normal frequency distribution; while the mode is 2-

point of the 2000 DSTP scores with a positively skewed frequency distribution. To examine the differences between the two score distributions from the field test and the 2000 DSTP, Chi-Square Goodness of Fit test, a non-parametric test, was conducted (See Table 4b). The results indicate that the relative frequency distributions of text-based writing scores from the two occasions are statistically significantly different at the $p < .00$ level for all the four grades.

Re-scoring Text-Based Writing Samples To further investigate the possible reasons that caused the low performance in text-based writing, a re-scoring session was arranged. A sample of 100 students' responses to the text-based task was randomly selected from the population of each grade. The panel members reviewed and re-scored the sample of writings independently. Each paper was evaluated holistically and analytically using the same 5-point scoring rubric. The analytic scores were only used as evidence to support their professional judgments on the quality of student writing. They worked in small groups to discuss their initial scores and assigned final scores (See Attachment F). Agreement among panel members was not required. Since the anchor papers were under review, the process of re-scoring was conducted without anchor papers. The records of re-scoring from the panel members can be found in Attachment G.

The different text-based writing scores (Diff = re-score - 2000 DSTP score) between the 2000 DSTP and re-scoring by the panel were calculated based on the following criteria:

- Majority rule was applied to decide the final re-score in the process of re-scoring.
- If the same number of readers agreed and the same number of readers did not agree with the 2000 DSTP score, the 2000 DSTP was used as the final re-score.
- If only one reader reviewed the paper, this reader's score was used as the final re-score.

Tables 5a and 5b present the summary of the discrepancies in scoring by grade. Data shows that the average agreement between the two scoring process is 65% across grades, 58% in grade 3, 66% in grade 5, 48% in grade 8, and 47% in grade 10 (Table 5b). Two patterns are observed: First, the scores assigned in the 2000 DSTP are consistently lower than that in re-scoring across the four grades; second, the distributions of the different scores are negatively skewed, which indicate more papers were scored lower than higher in the 2000 DSTP. On the average, 35% (113 papers) of the papers were scored 1-point lower in the 2000 DSTP than that in re-scoring, 8% (27 papers) of the papers were scored 2-points lower, and less than 1% (1 paper) of the papers were 3-points lower than they should be according to the panel's judgment. There are 9% (28 papers) of the papers that were scored 1-point higher and less than 1% (2 papers) of the papers scored 2-points higher than they should be across grades according to re-scoring. Grade by grade, it is found that 37% of the papers were scored lower and 5% of the papers scored higher in grade 3 from the 2000 DSTP; 29% of the papers were scored lower and 5% of the papers were scored higher based on re-scoring in grade 5. The data show that 46% of the papers were scored lower and 5% of the papers were scored higher in grade 8; 37% of the papers were scored lower and 16% of the papers were scored higher in the 2000 DSTP.

Tables 6a and 6b list the mean and frequency distributions of text-based scores from three occasions, field test, 2000 DSTP, and re-scoring. The data indicate that the writing scores re-assigned in this study are, on the average, higher than the 2000 DSTP, but lower than the field test in grades 3, 5, and 10. In both grades 3 and 10, the average re-score is closer to the 2000 DSTP in absolute value (.27 for grade 3; .16 for grade 10) than to the field test results (.76 for grade 3; .49 for grade 10). In grade 8, however, the average re-score is .38-point higher than the field test and .64-point (about one standard deviation) higher than the 2000 DSTP.

To examine how reliable the re-scoring process was, the number of readers used to re-score each paper and the degree of agreement among readers were analyzed. The majority of the papers were reviewed and re-scored by at least 3 readers, 96% in grade 3, 71% in grade 5, 51% in grade 8, and 68% in grade 10. In grades 3, 5, and 8, less than 5% of the papers were re-scored by a single reader; in grade 10, 22% of the papers were re-scored by one reader, which may be due to the content and the length of student writings in higher grade. In addition, the number of readers who assigned the same score that differs from the 2000 DSTP scores is used as the indicator of reliability of scoring (Table 7). The data indicate that the majority of the different scores were assigned by the panel in this study based on the agreement among 3 to 5 readers, 95% in grade 3, 70% in grade 5, 61% in grade 8, and 69% in grade 10. Fifteen percent of the papers (26 papers) receiving a different score were based on one reader; but most of them are 1-point higher rather than lower than the 2000 DSTP scores.

The panel members were also asked to make professional judgments on the difficulty level of the 1999 and 2000 text-based writing tasks by reviewing the reading passage and attached items. The results are summarized in Attachment F.

Text-Based Writing Development During group discussion, teachers provided comments and suggestions related to the development of the text-based writing. Their comments focused on three major issues: passage selection, wording of the prompt, and use of the writing rubric (See Attachment I).

- Passage Selection: Passages should be engaging and the difficulty level should be consistent from year to year. Third grade teachers preferred realistic stories as the basis for the text-based writing. Fifth grade teachers thought passages should be informative selections dealing with social studies or science.

- Wording of the Prompt: The wording in the prompt should always direct the students back to the text so that information from the text is included in the response. "Use details from the text to support your answer," should be in all prompts. Students should understand the concepts implied in the wording of the prompt. Developers should take care in using "user accessible" language in writing the prompts.

- Use of the Writing Rubric: Teachers discussed the possibility of adapting the general writing rubric so that each text-based writing item would have an item specific writing rubric. With specific guidelines provided in these item specific rubrics, it would be easier for scorers to determine when a student's writing is off topic.

Text-based Writing Instruction Teachers' comments related to instructional issues of the text-based writing focused on professional development in crafting text-based prompts and on identifying a variety of reading passages with which to write such prompts. They emphasized the need to have students write in response to a variety of text types (literary, informative, and technical) across content areas, and for teachers to model the process of making connections to the text and pulling out relevant details (See Attachment I).

- Tenth grade teachers pointed out that most of the writing done by high school students is text-based, and that text-based writing is not a separate *type* of writing. Written responses to texts are produced as *forms* of persuasive, expressive, or informative writing. High school teachers also expressed a concern regarding block scheduling, where students may have only five weeks of instruction prior to the administration of the DSTP. Finally, tenth grade teachers suggested that high school English teachers have gone away from literary analysis in lieu of an emphasis on stand-alone writing prompts, which may sacrifice students' writing in response to text.

- Fifth grade teachers stressed the importance of students making connections with characters in a story. They suggested that grade-level teams or district committees (lead by reading cadre representatives) develop questioning activities for teachers to use to improve students' performance on text-based writing. They also pointed out the need to release sample student responses to text-based writing prompts.

Limitations of the Study

As indicated in the beginning of this report, the current study was designed and conducted based on the available data within a short period of time. Due to the limitations of the study, the author suggests cautions in reviewing, interpreting, and using the results of this study.

- Information, such as sampling procedures and students' scores on the field test, is not available for review and additional analysis.
- Even though the sample of student text-based writings was randomly selected, the small sample size, only 1% of the grade population used in re-scoring, may not accurately reflect the characteristics of the population because of sampling errors. In addition, since the anchor papers were under review, the re-scoring process was conducted without using anchor papers.
- It is very important to note that previous studies have shown that the generalization of writing performance is low across the purpose (or discourse) of

writing tasks, writing topics, and occasions, especially when there are only a couple of items used in the writing assessment. In 2000, a new text-based writing task was introduced at each grade level, which may be one of the reasons for the fluctuation of the test scores. For example, third graders responded to an informative writing task instead of a persuasive writing task. Similarly, the fifth graders responded to an informative writing task in 1998 and 1999, but an expressive writing task in 2000. These changes could account for the low performance.

- An important issue of educational measurement is reliability. Reliability of performance-based assessment, such as writing, is often defined by agreement of readers in scoring a single task given on a single occasion, called inter-reader reliability or inter-reader consistency. However, another component of reliability involves the consistency of measurement over repeated occasions given fixed readers is called score reliability. Findings from early studies suggest that reader consistency differed considerably ranging from .33 to .91 and score reliability ranged from .26 to .60, which was dependent upon the number of points on the scoring scale, rating conditions, and changes in assessment programs (Dunbar et al, 1991; Fitzpatrick et al, 1994). The results of an experimental study conducted in Virginia (Moon et al, 1996) indicate that methods used for training and scoring (i.e., training readers to score multiple writing prompts at a single session or training readers sequentially score a new writing prompt) impact both reliability and validity. They also found that readers scored differently using the same scoring method on the same set of students' papers across years. To better understand the nature of direct writing assessment and provide valid and reliable measures of student achievement, more research questions, such as the stability of scoring over time, the process of reader training and scoring, and score reliability across topics, discourses, and occasions, need to be further explored.

References

Moon, T. R., Loyd, B. H., Highes, K. R. & Winter, P. (1996). Scoring and training issues involved in large-scale performance assessment.

Cantor, N. K. & Hoover, H. D. (1986). The reliability and validity of writing assessment: An investigation of rater, prompt within mode, and prompt between mode sources of error. Paper presented at the AERA Annual Conference, San Francisco, CA

Fitzpatrick, A. R., Ercikan, K., Yen, W. M. & Ferrara, S. (1994). The consistency between ratings collected in different test years. Paper presented at the NCME Annual Conference, New Orleans. LA

Dunbar, S. B., Kortez, D. M. & Hoover, H. D. (1991). Quality control in the development and use of performance assessment. Applied Measurement in Education. 4, 289-303.

State Summary Report -- Reading, Mathematics, Writing -- Spring 2000 Administration

State Summary Report in Writing -- 1998 DSTP

State Summary Report in Writing -- 1999 DSTP

State Summary Report in Writing -- 2000 DSTP

Table 1a

Means, Standard Deviations, and Frquency Distributions of Writing Scores by Grade and Year

GRADE 3 Writing Score	2000						1999						1998					
	Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
15																		
14					2	0.0											1	0.0
13					3	0.0					11	0.1					3	0.0
12					12	0.1					39	0.5					12	0.2
11					31	0.4					93	1.2					53	0.7
10			4	0.0	116	1.5			5	0.1	269	3.5			6	0.1	152	2.0
9			5	0.0	359	4.7			17	0.2	719	9.4			36	0.5	463	6.2
8			88	1.1	911	12.1			61	0.8	1137	14.8			138	1.8	928	12.4
7			261	3.4	1337	17.7			262	3.4	1477	19.2			438	5.8	1190	15.8
6			1447	19.2	1735	23.0			1149	14.9	1494	19.4			1740	23.2	1373	18.3
5	2	0.02	1729	23.0	1938	25.7	49	0.6	1801	23.4	1094	14.2	7	0.1	1523	20.3	1398	18.6
4	47	0.6	2737	36.4	643	8.5	387	5.0	2199	28.6	779	10.1	500	6.7	2353	31.3	1053	14.0
3	550	7.31	762	10.1	430	5.7	2147	27.9	1170	15.2	576	7.5	2132	28.4	739	9.8	520	6.9
2	2641	35.1	4.84	6.4			3391	44.1	1024	13.3			2847	37.9	541	7.2	367	4.9
1	4277	56.9					1714	22.3					2028	27.0				
Total	7517	100.0	7517	100.0	7517	100.0	7688	100.0	7688	100.0	7688	100.0	7514	100.0	7514	100.0	7514	100.0
Mean	1.52		4.54		6.06		2.18		4.26		6.44		2.15		4.70		6.85	
SD	0.66		1.27		1.63		0.85		1.40		1.94		0.90		1.41		1.98	
Purpose of Writing Task	Informative		Expressive				Persuasive		Informative				Persuasive		Expressive			

Table 1b

Means, Standard Deviations, and Frquency Distributions of Writing Scores by Grade and Year

GRADE 5	2000						1999						1998					
Writing	Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total	
Score	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
15					1	0.0											2	0
14					3	0.0					4	0.1					7	0.1
13					5	0.0					9	0.1					25	0.3
12					22	0.3					108	1.4					133	1.8
11					88	1.2					281	3.7					282	3.8
10			10	0.1	250	3.5			9	0.1	622	8.2			14	0.2	678	9.2
9			6	0.0	718	10.1			5	0.1	1402	18.4			32	0.4	1065	14.4
8			121	1.7	1405	19.8			349	4.6	1453	19.1			247	3.3	1208	16.3
7			398	5.6	1477	20.8			498	6.5	1294	17.0			518	7.0	1473	19.9
6			2424	34.2	1342	18.9			2573	33.8	1344	17.7			1682	22.7	1325	17.9
5	8	0.0	1528	21.5	1189	16.7	4	0.1	1384	18.2	794	10.4	8	0.1	1462	19.8	769	10.4
4	155	2.2	1898	26.7	374	5.2	624	8.2	2350	30.9	192	2.5	1205	16.3	2756	37.3	249	3.4
3	974	13.7	446	6.2	210	2.9	2761	36.3	278	3.7	105	1.4	2596	35.1	396	5.4	178	2.4
2	2958	41.8	252	3.5			2850	37.5	162	2.1			2365	32.0	287	3.9		
1	2988	42.2					1369	18					1220	16.5				
Total	7084	100.0	7084	100.0	7084	100.0	7608	100.0	7608	100.0	7608	100.0	7394	100.0	7394	100.0	7394	100.0
Mean	1.76		5.02		6.78		2.35		5.17		7.52		2.52		4.89		7.41	
SD	0.78		1.27		1.74		0.87		1.29		1.86		0.96		1.36		1.98	
Purpose of Writing Task	Expressive		Persuasive				Informative		Informative				Informative		Persuasive			

Table 1c

Means, Standard Deviations, and Frquency Distributions of Writing Scores by Grade and Year

GRADE 8	2000						1999						1998					
Writing	Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total	
Score	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
15					1	0.0					1	0.0						
14					1	0.0					5	0.1					10	0.1
13					6	0.0					10	0.1					33	0.4
12					22	0.2					78	1.0					205	2.6
11					102	1.3					184	2.4					497	6.3
10			7	0.0	353	4.5			8	0.1	572	7.3			208	0.2	763	9.9
9			5	0.0	948	12.3			25	0.3	1230	15.8			353	0.9	1309	16.5
8			220	2.8	2483	32.2			183	2.3	1757	22.5			1710	7.6	1514	19.1
7			761	9.8	1795	23.3			764	9.8	1531	19.6			1566	12.7	1266	16.0
6			3691	47.9	1070	13.9			2755	35.3	1228	15.8			2393	30.2	1152	14.5
5	1	0.01	1633	21.2	770	10.0	5	0.1	2006	25.7	825	10.6	2	0.0	1011	19.7	757	9.5
4	76	0.9	1204	15.6	108	1.4	360	4.6	1589	20.4	243	3.1	593	7.5	606	21.6	239	3.0
3	902	11.7	135	1.7	26	0.3	2069	26.5	303	3.9	130	1.7	2519	31.7	71	4.4	169	2.1
2	4254	55.3	34	0.4			3297	42.3	161	2.1			3255	41.0	16	2.6		
1	2457	31.9					2063	26.5					1565	19.7				
Total	7685	100.0	7685	100.0	7685	100.0	7794	100.0	7794	100.0	7794	100.0	7934	100.0	7934	100.0	7934	100.0
Mean	1.82		5.57		7.39		2.10		5.29		7.39		2.27		5.45		7.72	
SD	0.67		1.05		1.46		0.84		1.22		1.79		0.86		1.46		2.06	
Purpose of Writing Task	Expressive		Informative				Persuasive		Informative				Persuasive		Informative			

Table 1d

Means, Standard Deviations, and Frquency Distributions of Writing Scores by Grade and Year

GRADE 10	2000						1999						1998					
Writing	Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total		Text-Based		Stand-Alone		Writing Total	
Score	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
15					1	0.0												
14					2	0.0												
13					9	0.1					11	0.2					1	0.0
12					76	1.1					29	0.4					16	0.2
11					189	2.7					170	2.5					113	1.7
10			6	0.0	427	6.1			422	0.1	362	5.4					321	4.8
9			10	0.1	970	14.0			568	0.2	780	11.6			1	0.0	848	12.6
8			267	3.8	1644	23.8			1948	3.1	1092	16.3			68	1.0	1185	17.6
7			628	9.1	1312	19.0			1314	8.6	1239	18.5			392	5.8	1248	18.6
6			2703	39.2	1126	16.3			1646	24.6	1265	18.9			1703	25.3	1508	22.4
5	10	0.1	1447	20.9	892	12.9			575	19.6	989	14.8			1426	21.2	1156	17.2
4	305	4.4	1548	22.4	156	2.2	115	1.7	207	29.1	453	6.8	295	4.4	2724	40.5	208	3.1
3	1268	18.3	183	2.6	90	1.3	1536	22.9	16	8.5	312	4.7	1687	25.1	271	4.0	120	1.8
2	3132	45.4	102	1.4			2903	43.3	6	6.3			2910	43.3	139	2.1		
1	2179	31.6					2148	32.1					1823	27.2				
Total	6894	100.0	6894	100.0	6894	100.0	6702	100.0	6702	100.0	6702	100.0	6724	100.0	6724	100.0	6724	100.0
Mean	1.96		5.38		7.34		1.94		4.87		6.82		2.07		4.85		6.92	
SD	0.83		1.21		1.77		0.78		1.45		1.95		0.83		1.15		1.71	
Purpose of Writing Task	Persuasive		Informative				Persuasive		Informative				Persuasive		Informative			

Table 2

Means and Standard Deviations of Writing Scores by Gender, Grade, and Year

Grade	Gender	Type of Writing	2000			1999			1998		
			N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
3	*female*	stand-alone	3747	4.73	1.26	3756	4.48	1.38	3708	4.98	1.39
		text-based	3747	1.58	0.69	3756	2.29	0.86	3708	2.30	0.91
		writing total	3747	6.32	1.63	3756	6.76	1.92	3708	7.29	1.96
	male	stand-alone	3770	4.35	1.25	3927	4.06	1.39	3804	4.43	1.38
		text-based	3770	1.45	0.62	3927	2.07	0.83	3804	2.00	0.86
		writing total	3770	5.81	1.58	3927	6.13	1.91	3804	6.43	1.91
5	*female*	stand-alone	3546	5.25	1.23	3717	5.41	1.28	3755	5.13	1.36
		text-based	3546	1.89	0.80	3717	2.49	0.87	3755	2.67	0.95
		writing total	3546	7.14	1.72	3717	7.91	1.83	3755	7.79	1.97
	male	stand-alone	3537	4.78	1.27	3886	4.94	1.26	3633	4.67	1.31
		text-based	3537	1.63	0.73	3886	2.21	0.84	3633	2.36	0.93
		writing total	3538	6.42	1.68	3886	7.15	1.81	3633	7.03	1.91
8	*female*	stand-alone	3874	5.71	1.02	3940	5.56	1.14	3889	5.78	1.40
		text-based	3874	1.93	0.69	3940	2.25	0.85	3889	2.45	0.85
		writing total	3874	7.65	1.44	3940	7.81	1.69	3889	8.23	1.96
	male	stand-alone	3810	5.42	1.07	3793	5.04	1.24	4045	5.13	1.44
		text-based	3810	1.70	0.62	3793	1.95	0.81	4045	2.10	0.84
		writing total	3811	7.12	1.44	3793	6.99	1.78	4045	7.23	2.02
10	*female*	stand-alone	3547	5.56	1.17	3399	5.10	1.40	3455	5.10	1.14
		text-based	3547	2.10	0.83	3399	2.06	0.78	3455	2.24	0.83
		writing total	3547	7.66	1.72	3399	7.16	1.87	3455	7.34	1.68
	male	stand-alone	3347	5.18	1.23	3263	4.66	1.46	3269	4.59	1.09
		text-based	3347	1.82	0.81	3263	1.83	0.77	3269	1.89	0.80
		writing total	3347	7.00	1.75	3263	6.49	1.96	3269	6.48	1.62

Table 3a

Correlation Matrix of DSTP Writing, DSTP Reading, and SAT9 Reading Scores by Grade

GRADE 3

	MC Items	CR Items	Passage	Item 67	Text-based	Stand-alone	Writing	Reading	SAT9
MC Items	1.00	0.46	0.77	0.24	0.19	0.27	0.29	0.63	0.51
CR Items		1.00	0.92	0.66	0.31	0.36	0.40	0.71	0.51
Passage			1.00	0.58	0.31	0.38	0.42	0.78	0.59
Item 67 – scored for ending writing				1.00	0.22	0.24	0.28	0.44	0.29
Text-based					1.00	0.36	0.69	0.37	0.28
Stand-alone						1.00	0.93	0.50	0.43
Writing							1.00	0.53	0.45
Reading								1.00	0.88
SAT9									1.00

GRADE 5

	MC Items	CR Items	Passage	Item 56	Text-based	Stand-alone	Writing	Reading	SAT9
MC Items	1.00	0.40	0.75	0.31	0.26	0.31	0.35	0.65	0.52
CR Items		1.00	0.91	0.73	0.44	0.42	0.50	0.66	0.45
Passage			1.00	0.67	0.44	0.45	0.52	0.77	0.56
Item 56				1.00	0.45	0.37	0.47	0.51	0.37
Text-based					1.00	0.41	0.75	0.44	0.33
Stand-alone						1.00	0.91	0.50	0.41
Writing							1.00	0.56	0.45
Reading								1.00	0.86
SAT9									1.00

Table 3b

Correlation Matrix of DSTP Writing, DSTP Reading, and SAT9 Reading Scores by Grade

GRADE 8

	MC Items	CR Items	Passage	Item 47	Text-based	Stand-alone	Writing	Reading	SAT9
MC Items	1.00	0.48	0.79	0.34	0.34	0.38	0.43	0.69	0.58
CR Items		1.00	0.92	0.64	0.46	0.41	0.50	0.71	0.52
Passage			1.00	0.61	0.47	0.46	0.54	0.81	0.63
Item 47				1.00	0.57	0.37	0.53	0.56	0.38
Text-based					1.00	0.41	0.75	0.52	0.39
Stand-alone						1.00	0.91	0.55	0.48
Writing							1.00	0.63	0.52
Reading								1.00	0.85
SAT9									1.00

GRADE 10

	MC Items	CR Items	Passage	Item 46	Text-based	Stand-alone	Writing	Reading	SAT9
MC Items	1.00	0.35	0.70	0.32	0.30	0.35	0.38	0.59	0.50
CR Items		1.00	0.91	0.70	0.47	0.44	0.52	0.64	0.44
Passage			1.00	0.67	0.49	0.49	0.56	0.74	0.55
Item 46				1.00	0.60	0.46	0.60	0.59	0.40
Text-based					1.00	0.48	0.80	0.54	0.39
Stand-alone						1.00	0.91	0.59	0.48
Writing							1.00	0.66	0.52
Reading								1.00	0.84
SAT9									1.00

Table 4

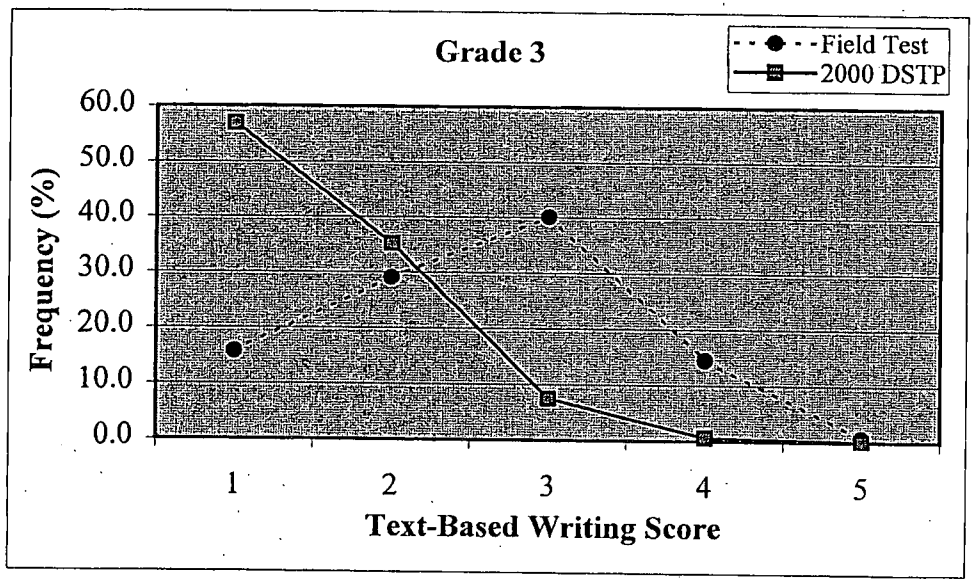
Means and Frequency Distributions of Writing Scores in Field Test and 2000 DSTP

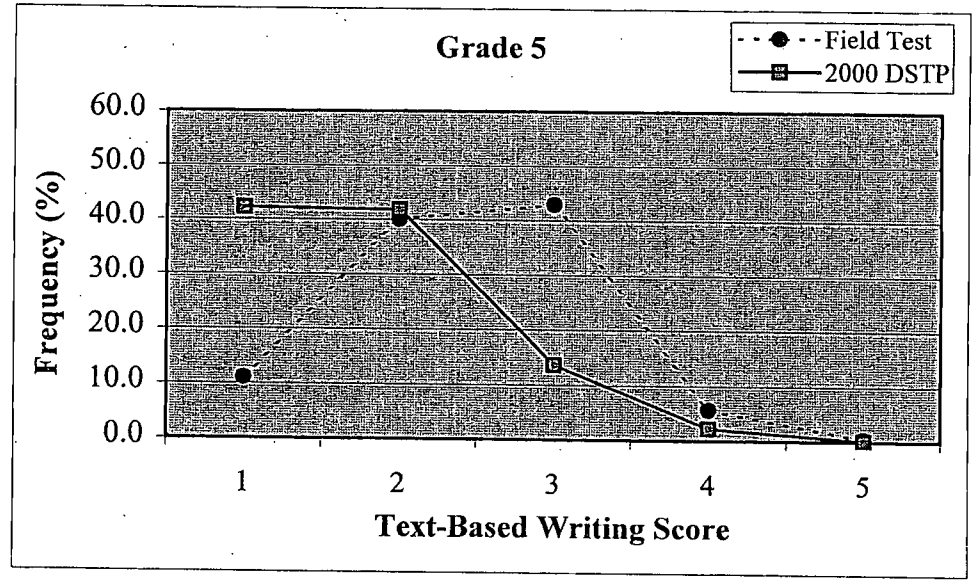
	Grade 3				Grade 5				Grade 8				Grade 10			
score	field test		2000 DSTP		field test		2000 DSTP		field test		2000 DSTP		field test		2000 DSTP	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
5	6	0.7	2	0.0	4	0.4	8	0.1	8	0.9	1	0.0	21	2.5	10	0.1
4	132	14.5	47	0.6	50	5.6	155	2.2	52	6.1	76	0.9	139	16.5	304	4.4
3	364	40.1	550	7.4	381	42.8	974	13.7	190	22.5	902	11.8	295	35.1	1268	18.4
2	263	29.0	2641	35.1	356	40.1	2958	41.8	351	41.5	4254	55.4	263	31.2	3132	45.5
1	143	15.7	4277	56.9	99	11.1	2988	42.2	246	29.0	2457	31.9	124	14.7	2179	31.6
mean	2.55	100.0	1.52	100.0	2.44	100.0	1.76	100.0	2.08	100.0	1.82	100.0	2.61	100.0	1.96	100.0
N.	908		7517		890		7084		847		7685		842		6894	

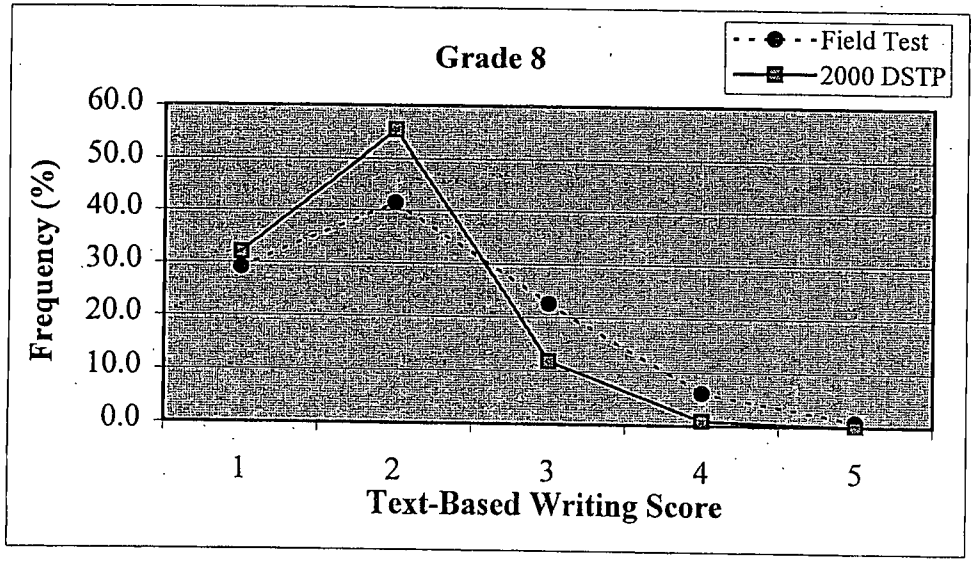
* The field test results are based on the writing scores for students in grades 4, 6, 9, and 11.

Chart 1

Comparisons of Percentage of Frequency Distributions of Text-Based Scores Between Field Test and 2000 DSTP







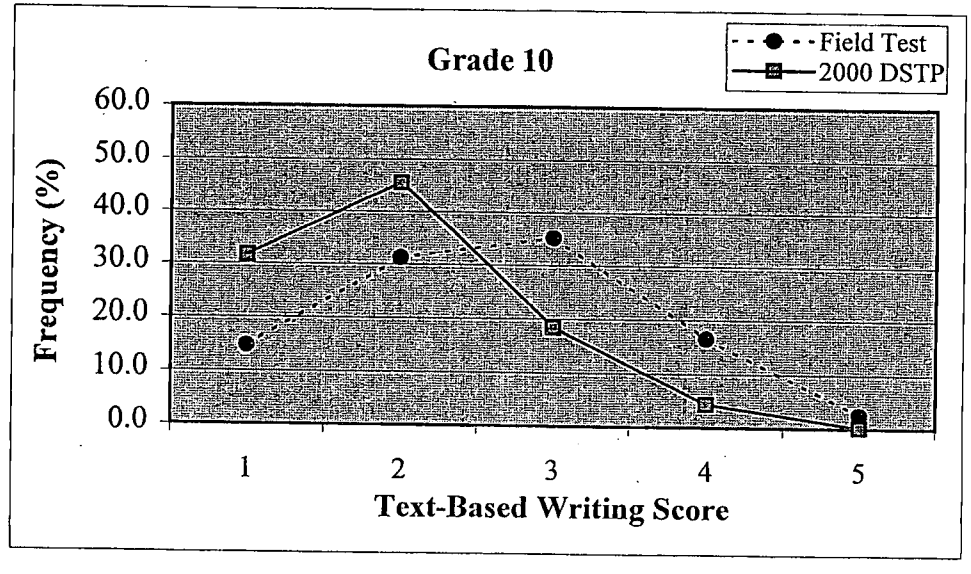


Table 5a

Summary of Discrepancies of Text-Based Writing Scores between 2000 DSTP and Re-Scoring

Diff Score	Grade 3		Grade 5		Grade 8		Grade 10		Sub-Total	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
3	0	0.0	0	0.0	1	1.0	0	0.0	1	0.3
2	10	10.0	4	4.0	7	7.0	6	6.0	27	6.8
1	25	25.0	24	24.0	36	36.0	28	28.0	113	28.3
0	55	55.0	65	65.0	45	45.0	44	44.0	209	52.3
-1	3	3.0	5	5.0	5	5.0	15	15.0	28	7.0
-2	2	2.0	0	0.0	0	0.0	0	0.0	2	0.5
-3	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
*	5	5.0	2	2.0	6	6.0	7	7.0	20	5.0
Sub-Total	100	100.0	100	100.0	100	100.0	100	100.0	400	100.0

Table 5b

Summary of Discrepancies of Text-Based Writing Scores between 2000 DSTP and Re-Scoring (Adjusted)

Diff Score	Grade 3		Grade 5		Grade 8		Grade 10		Sub-Total	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
3	0	0.0	0	0.0	1	1.1	0	0.0	1	0.3
2	10	10.5	4	4.1	7	7.4	6	6.5	27	8.4
1	25	26.3	24	24.5	36	38.3	28	30.1	113	35.0
0	55	57.9	65	66.3	45	47.9	44	47.3	209	64.7
-1	3	3.2	5	5.1	5	5.3	15	16.1	28	8.7
-2	2	2.1	0	0.0	0	0.0	0	0.0	2	0.6
-3	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Sub-Total	95	100.0	98	100.0	94	100.0	93	100.0	380	117.6

Diff Score = Re-Score - Original Score

* = Number of papers were non-scorable because of missing pages.

Table 6a

Summary of Discrepancies in Scoring Text-Based Writing from field test, 2000 DSTP, and Re-Scoring

score	Grade 3						Grade 5					
	field test		2000 DSTP		re-scoring		field test		2000 DSTP		re-scoring	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
5	6	0.7	2	0.0	0	0.0	4	0.4	8	0.1	1	1.0
4	132	14.5	47	0.6	1	1.1	50	5.6	155	2.2	2	2.1
3	364	40.1	550	7.4	17	17.9	381	42.8	974	13.7	19	19.6
2	263	29.0	2641	35.1	39	41.1	356	40.1	2958	41.8	42	43.3
1	143	15.7	4277	56.9	38	40.0	99	11.1	2988	42.2	33	34.0
mean	2.55	100.0	1.52	100.0	*1.79*	100.0	2.44	100.0	1.76	100.0	*2.09*	100.0
N.	908		7517		95		890		7084		97	

Table 6b

Summary of Discrepancies of Text-Based Writing Scores in Field Test, 2000 DSTP, and Re-Scoring

score	Grade 8						Grade 10					
	field test		2000 DSTP		re-scoring		field test		2000 DSTP		re-scoring	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
5	8	0.9	1	0.0	0	0.0	21	2.5	10	0.1	0	0.0
4	52	6.1	76	0.9	8	8.5	139	16.5	304	4.4	6	6.5
3	190	22.5	902	11.8	28	29.8	295	35.1	1268	18.4	32	34.4
2	351	41.5	4254	55.4	44	46.8	263	31.2	3132	45.5	34	36.6
1	246	29.0	2457	31.9	14	14.9	124	14.7	2179	31.6	18	19.4
mean	2.08	100.0	1.82	100.0	*2.46*	100.0	2.61	100.0	1.96	100.0	*2.12*	96.8
N.	847		7685		94		842		6894		90	

Zero scores are not included from re-scoring for equivalent comparison with the field test and the 2000 DSTP.

Table 7

Number of Readers Used in Re-Scoring Text-Based Writing

N. of Readers	Grade 3		Grade 5		Grade 8		Grade 10	
	N.	%	N.	%	N.	%	N.	%
1	2	2.1	1	1.0	5	5.3	20	21.5
2	2	2.1	27	27.6	41	43.6	10	10.8
3	10	10.5	37	37.8	38	40.4	34	36.6
4	22	23.2	26	26.5	6	6.4	20	21.5
5	59	62.1	7	7.1	4	4.3	9	9.7
Sub-Total	95	100.0	98	100.0	94	100.0	93	100.0

The data presented in this table show the number of readers used in re-scoring the text-based wriitng in this study.

Table 8

Reliability Indicator of Text-Based Writing Scores in Re-Scoring

Diff Score	Grade 3 Number of Readers					Grade 5 Number of Readers					Grade 8 Number of Readers					Grade 10 Number of Readers				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
3												1								
2			1	4	5		2	1	1			2	3	1	1	1		4		1
1		1	3	3	18	9	9	5	1		2	14	17	2	1	9	1	7	10	1
-1		1			2	1	2	1	1				3	1	1	2	2	7	3	1
-2					2															
-3																				
Total		2	4	7	27	10	13	7	3		2	17	23	4	3	12	3	18	13	3
%		5.0	10.0	17.5	67.5	30.3	39.4	21.2	9.1		4.1	34.7	46.9	8.2	6.1	24.5	6.1	36.7	26.5	6.1

The data shown in this table only include the different text-based writing scores between 2000 DSTP and re-scoring.

Attachment A
Writing Scoring Rubric

Delaware Student Testing Program - General Rubric for Writing

The following characteristics determine the success of the response in meeting the needs of the audience and fulfilling the writing purpose.

Score of 5	Score of 4	Score of 3	Score of 2	Score of 1
Score point 5 meets all the criteria listed in score point 4. In addition, a paper receiving this score shows an exceptional awareness of readers' concerns and needs. The student may have shown an exceptional use of: • Development strategies specific to the purpose for writing • Distinctive style, voice, tone • Literary devices • Compositional risks	Unified with smooth transitions, a clear and logical progression of ideas, and an effective introduction and closing. Sufficient, specific, and relevant details that are fully elaborated. Consistently complete sentences with appropriate variety in length and structure. A consistent style with precise and vivid word choice. Few, if any, errors in standard written English that do not interfere with understanding.	Generally unified with some transitions, a clear progression of ideas, and an introduction and closing. Specific details but may be insufficient, irrelevant, or not fully elaborated. Generally complete sentences with sufficient variety in length and structure. Some style and generally precise word choice. Some errors in standard written English that rarely interfere with understanding.	Minimally unified and may lack transitions or an introduction or closing. Some specific details but may be insufficient, irrelevant, and/or not elaborated. Some sentence formation errors and a lack of sentence variety. Sometimes general and repetitive word choice. Several kinds of errors in standard written English that interfere with understanding.	Lacks unity. No or few specific details that are minimally elaborated. Frequent and severe sentence formation errors and/or a lack of sentence variety. Often general, repetitive, and/or confusing word choice. Frequent and severe errors in standard written English that interfere with understanding.

For non-scorable responses see below:

- Blank
- Off topic
- Written in a language other than English
- Refusal
- Illegible
- Insufficient

/elp
5/10/99

Attachment B
Instructions for Anchor Paper Review

INSTRUCTIONS FOR ANCHOR PAPER REVIEW

1. In the folder, you will find the following:

   - Two copies of the Form of Anchor Paper Review, one for each grade
   - Non-disclosure Form
   - Writing Scoring Rubric
   - Reading passage and the question for the text-based writing
   - Anchor papers for selected grades

2. Procedures for Anchor Paper Review:

Step One: (On your own)

- Sign the Non-disclosure Form.
- Read the Instructions for Anchor Paper Review.
- Read the reading passage and the question for the text-based writing.
- Read the scoring rubric.

Step Two: (On your own)

- Read each anchor paper carefully.
- Rank all the anchor papers from the lowest (as #1) to the highest (whatever the number applies). You may assign the same rank order to more than one paper, such as two papers are ranked as #3 and the following one should be #5 instead of #4.
- Assign a score to each anchor paper using the 5-point rubric. You may use high or low to differentiate papers even though they receive the same score point(s).
- Record your scores using the Form of Anchor Paper Review. Match the Code with the 1-digit number sequential number on the right top of each paper and fill in the rank order and the score you have assigned to each paper.
- You also need to record the 11-digit paper number (located on the right bottom of each paper) in under the column of Paper Number of the Form of Anchor Paper Review.
- Provide your comments using the space below the Form.
- Please return the entire folder to Kim by August 7, 2000.

Step Three: (Group meeting if necessary)

- Have a group meeting to share your scores with other reviewers.
- Group Discussion
- Make your second judgment
- Compare your judgments with the original scores.

Attachment C
Procedures For Re-scoring

Agenda

Study for the DSTP Writing Assessment
8:30 am-3:30 pm, Tatnall Building, Dover, August 17, 2000

Welcome and Study Brief (Zhang)	8:30-8:45
Scoring Rubrics Training	8:45-9:05
• *Holistic Scoring Rubrics for Writing (Bolig)* • *Holistic Scoring Rubrics for Reading (Bolig)* • *Analytic Scoring Rubrics for Writing (Kelley)*	
Evaluate Text-based Writing Tasks	9:05-9:30
• *First Round Evaluation*	
Review and Re-score Students' Writings	9:30-11:00
• *First Round Scoring*	
Small Group Discussion	11:00-12:00
Lunch	12:00-12:40
Continue Group Discussion and Second Round Scoring	12:40-2:00
Evaluate Text-based Writing Tasks	2:00-2:30
• *Second Round Scoring*	
Grade-Group Discussion	2:30-3:30
• *Recommendations* • *Comments* • *Evaluation*	

Appendix A

Delaware English Language Arts Content Standards

Students in Delaware public schools, using the process of effective readers, writers, listeners, viewers, and speakers, will be able to:

Standard #1
Use written and oral English appropriate for various purposes and audiences.

Standard #2
Construct, examine, and extend the meaning of literacy, informative, and technical texts through listening, reading, and writing.

Standard #3
Access, organize, and evaluate information gained through listening, reading, and viewing.

Standard #4
Use literary knowledge accessed through print and visual media to connect self to society and culture.

Appendix B1

Definitions of Cognitive Categories for DSTP Reading Comprehension Assessment

Three stances of are used to measure the depth of reading comprehension. These stances are: Determining Meaning, Interpreting Meaning, and Extending Meaning.

- Questions in the Determining Meaning stance require the reader to demonstrate an overall understanding of the passage. The focus is on how the reader begins to make meaning of the text.
- Questions in the Interpreting Meaning stance require the reader to go beyond the initial understanding to develop an interpretation of the text. The reader goes beyond first impression to construct a more complete understanding of what has been read.
- Questions in the Extending Meaning stance require the reader to stand apart from the text and critically consider it. This stance involves critical examination, evaluation, and analysis.

Delaware Student Testing Program – Instructional Guide for Writing

(An Analytic Adaptation of the DSTP General Rubric for Writing)

The following characteristics determine the success of the response in meeting the needs of the audience and fulfilling the writing purpose.

Score of 5		Score of 4	Score of 3	Score of 2	Score of 1
Score point 5 meets all the criteria listed in score point 4. In addition, a paper receiving this score shows an exceptional awareness of readers' concerns and needs. *The student may have shown an exceptional use of:* • *Development strategies specific to the purpose for writing* • *Distinctive style, voice, tone* • *Literary devices* • *Compositional risks*	Organization	Unified with smooth transitions, a clear and logical progression of ideas, and an effective introduction and closing.	Generally unified with some transitions, a clear progression of ideas, and an introduction and closing.	Minimally unified and may lack transitions or an introduction or closing.	Lacks unity.
	Development	Sufficient, specific, and relevant details that are fully elaborated.	Specific details but may be insufficient, irrelevant, or not fully elaborated.	Some specific details but may be insufficient, irrelevant, and/or not elaborated.	No or few specific details that are minimally elaborated.
	Sentence Formation	Consistently complete sentences with appropriate variety in length and structure.	Generally complete sentences with sufficient variety in length and structure.	Some sentence formation errors and a lack of sentence variety.	Frequent and severe sentence formation errors and/or a lack of sentence variety.
	Style/Word Choice	A consistent style with precise and vivid word choice.	Some style and generally precise word choice.	Sometimes general and repetitive word choice.	Often general, repetitive, and/or confusing word choice.
	Conventions	Few, if any, errors in standard written English that do not interfere with understanding.	Some errors in standard written English that rarely interfere with understanding.	Several kinds of errors in standard written English that interfere with understanding.	Frequent and severe errors in standard written English that interfere with understanding.

The Joy of Scoring

- One part holistic
- Five parts analytic
- Apply to writing samples
- Pray for inter-rater reliability

Holistic vs. Analytic Scores

There is no concrete correlation between the holistic score for a piece of writing and the analytic scores for each of the textual features at work within that piece of writing. A holistic 3 does not mean that each analytic score is also a 3. Nor does a holistic 3 mean that the sum of the analytic scores is 15, which when divided by 5 (the number of textual features) would "give us" that holistic 3. It's not that simple.

First, the textual features are not weighted equally. Generally, Organization and Development carry far more clout than Conventions. Sentence Formation and Style/Word Choice, while also considered more significant than Conventions, are not quite as critical to a piece as Organization and Development. Additionally, the nature of a particular prompt, its audience, or its purpose may subtly emphasize a particular textual feature and thereby influence the scoring.

Second, in addition to the distinct analytic scores, a holistic score also takes into account how those individual textual features *work together* to create the *whole* piece of writing. In essence, the whole is "greater" than the sum of its parts. But "greater" does not imply that the holistic score is higher; it means that there is more to consider in the holistic score than just the analytic scores for each of the textual features.

Our Task

When analytically scoring the samples for this study, consider only the four levels below. Score points 4 and 5 from the DSTP Rubric are lumped together for the purposes of this study. We are less concerned with determining the differences between those score points than we are with determining the differences, in terms of the textual features, between score points 3, 2, and 1.

Score Point from DSTP Rubric	4/5	3	2	1
Descriptor	Very Good	Good Enough	So-So	Not So Hot

The Chicken and the Egg: Which comes first, holistic or analytic?

What holistic score should we assign to a paper with the following configuration of analytic scores? Why?

Example A

Textual Feature	Analytic Score
Organization	1/2
Development	1/2
Sentence Formation	3/4
Style/Word Choice	3/4
Conventions	3/4

Example B

Textual Feature	Analytic Score
Organization	1/2
Development	2/3
Sentence Formation	3/4
Style/Word Choice	3/4
Conventions	3/4

Example C

Textual Feature	Analytic Score
Organization	1/2
Development	1/2
Sentence Formation	2/3
Style/Word Choice	2/3
Conventions	3/4

Example D

Textual Feature	Analytic Score
Organization	3/4
Development	3/4
Sentence Formation	3/4
Style/Word Choice	3/4
Conventions	3/4

Example E

Textual Feature	Analytic Score
Organization	2/3
Development	3/4
Sentence Formation	2/3
Style/Word Choice	2/3
Conventions	3/4

Example F

Textual Feature	Analytic Score
Organization	3/4
Development	3/4
Sentence Formation	2/3
Style/Word Choice	2/3
Conventions	3/4

- Begin developing a sense of the holistic score by looking at Organization and Development. These two textual features carry the most weight.
- The combination of analytic scores for Sentence Formation, Style/Word Choice, and Conventions may shift—either up or down--the holistic score developed by considering the analytic scores for Organization and Development. This shift will seldom be more than one score point.
- Conventions alone do not make a piece of writing *better*; they merely make it *correct* in terms of Standard Written English. Conventional correctness does not improve the *quality* of a piece of writing that lacks unity, specific detail, sentence variety, style, etc. Quality comes with improvements in the other textual features. Conventions do make the piece more readable, but readable writing that says nothing and makes no sense still says nothing and makes no sense.

Attachment D
List of Panel Members

Anchor Paper Review Committee
For FY 2000 DSTP Writing Study

Name	Gender	District	School	Test Development Committee
Grade 3 & 5				
Cookie Bolig	F	Department of Ed.	Assessment & Analysis	Y
Jackie Shockley	F	Cape Henlopen	R.A. Shields Elementary	Y
Marty Hodgkins	F	Appoquinimink	Redding Intermediate	Y
Linda Mitchell	F	Indian River	East Millsboro Elementary	N
Chris Evans	F	Brandywine	Mt. Pleasant High	Y
Grade 8 & 10				
Mike Boyd	M	Lake Forest	Lake Forest High	Y
Kate Szegda	F	Red Clay	District Office	Y
Doug Grudzina	M	Capital	Dover High	N
Mike Kelly	M	Department of Ed.	Curriculum Development	Y

Anchor Paper Pulling Review Committee
For DSTP Writing Study

Name	Gender	District	School	Test Development Committee
Grade 3				
Jackie Shockley	F	Cape Henlopen	Shields Elementary	Y
Janice Trainer	F	Christina	Wilson Elementary	N
Ann Whitman	F	Milford	Benjamin Banneker Elementary	N
Nancy Carter	F	Milford	Benjamin Banneker Elementary	N
Cookie Bolig	F	Department of Ed.	Assessment & Analysis	Y
Grade 5				
Linda Mitchell	F	Indian River	East Millsboro	N
Marty Hodgkins	F	Appoquinimink	Redding Intermediate	Y
Mary Currie	F	Milford	Milford Middle	Y
Christine Poehlmann	F	Appoquinimink	Redding Intermediate	N
Karen Sheets	F	Woodbridge	Woodbridge Elementary	N
Grade 8				
Kate Szegda	F	Red Clay	District Office	Y
Ginger Angstadt	F	Capital	Central Middle School	Y
Linda Poorman	F	Colonial	Wallace Wallin	N
Gwnne Ash	F	University of DE		N
Denise Speicher	F	Indian River	District Office	Y
Grade 10				
Mike Boyd	M	Lake Forest	Lake Forest High	Y
Catherine Laverick	F	NCCVT	Hodgson Vo-Tech	N
John Drumheller	M	Cape Henlopen	Cape Henlopen High	N
Judy Smith	F	Polytech	Polytech High	N
Pam Wilson	F	Polytech	Polytech High	N
Doug Grudzina	M	Capital	Dover High	N
Mike Kelly	M	Department of Ed.	Curriculum Development	Y

Attachment E
Correlation Matrix
Between Reading and Writing

Correlation Matrix between Reading and Writing

1999	READING	TEXT-BASED	STAND-ALONE	WRITING
Grade 3				
READING	*1.00*			
TEXT-BASED	*0.60*	*1.00*		
STAND-ALONE	*0.57*	*0.45*	*1.00*	
WRITING TOTAL	*0.68*	*0.77*	*0.92*	*1.00*
Grade 5				
Reading	*1.00*			
Text-Based	*0.56*	*1.00*		
Prompt	*0.55*	*0.46*	*1.00*	
Writing	*0.64*	*0.79*	*0.91*	*1.00*
Grade 8				
Reading	*1.00*			
Text-Based	*0.65*	*1.00*		
Prompt	*0.55*	*0.49*	*1.00*	
Writing	*0.68*	*0.80*	*0.91*	*1.00*
Grade 10				
Reading	*1.00*			
Text-Based	*0.55*	*1.00*		
Prompt	*0.59*	*0.47*	*1.00*	
Writing	*0.66*	*0.75*	*0.92*	*1.00*
1998	Reading	Text-B	Prompt	Writing
Grade 3				
Reading	*1.00*			
Text-Based	*0.56*	*1.00*		
Prompt	*0.52*	*0.44*	*1.00*	
Writing	*0.63*	*0.77*	*0.91*	*1.00*
Grade 5				
Reading	*1.00*			
Text-Based	*0.60*	*1.00*		
Prompt	*0.57*	*0.45*	*1.00*	
Writing	*0.68*	*0.79*	*0.90*	*1.00*
Grade 8				
Reading	*1.00*			
Text-Based	*0.60*	*1.00*		
Prompt	*0.63*	*0.54*	*1.00*	
Writing	*0.70*	*0.80*	*0.94*	*1.00*
Grade 10				
Reading	*1.00*			
Text-Based	*0.56*	*1.00*		
Prompt	*0.56*	*0.47*	*1.00*	
Writing	*0.65*	*0.81*	*0.90*	*1.00*

* All correlation coefficients are calculated based on aggregated data.

Attachment F
Summary of Anchor Paper Review

Summary of Anchor Paper Review
Grade 3

Sequential		*1st Round*						*2nd Round*						*Orig.*	*Diff*	*Notes*
Number	**Paper N.**	**1**	**2**	**3**	**4**	**5**		**1**	**2**	**3**	**4**	**5**	mean	**score**	**score**	
1	57611004430	1	2-	1	1	1		1	1.5	1	1	1	1.1	1	0.1	#1 & #2 - Long sentence! Probably scored better for reading.
2	57863009470	1	1	1	1	1		1	1	1	1	1	1	1	0	
3	57854000210	1+	1	1+	1	1		1.5	1	1.5	1	1	1.2	1	0.2	#3 - Point of view few details.
4	57863010700	2-	2	2-	1	2		1.5	2	1.5	2	2	1.8	2	-0.2	
5	58468003550	2	2+	2	2	2		2	2.5	2	2	2	2.1	2	0.1	
6	57854011760	2	2-	2	2	2		2	1.5	2	2	2	1.9	2	-0.1	#6 - Should score low for reading.
7	57854000140	3	3	3-	3	3-		3	3	2.5	3	2.5	2.8	2	**0.8**	
8	58109020110	3	3	3	2	3-		3	3	3	2	2.5	2.7	3	-0.3	
9	58270009250	3+	3+	3+	3	3		3.5	3.5	3.5	3	3	3.3	3	0.3	
10	58874001980	3+	4+	4	3	3		3.5	4.5	4	3	3	3.6	3.5	0.1	#10 - Original score should be a 3+ according to the rubric. It meets the aspects of a 3. It does not meet all aspects consistently for a 4.
11	58109028280	4-	3+	4+	3	3+		3.5	3.5	4.5	3	3.5	3.6	4	-0.4	
12	58280011260	4+	4	5	4	4-		4.5	4	5	4	3.5	4.2	4	0.2	
Discrepancy															1	
%															8%	

Comments

Summary of An...or Paper Review
Grade 5

Sequential Number	Paper N.	*1st Round*						*2nd Round*						*Orig.*	*Diff*	*Notes*
		1	2	3	4	5		1	2	3	4	5	mean	score	score	
1	58083017170	1	1	1	1	1		1	1	1	1	1	1	0.5	0.5	
2	58083010620	1	1	1	1	1		1	1	1	1	1	1	1	0	
3	57854001820	1+	2-	1	1	1		1.5	1.5	1	1	1	1.2	1.5	-0.3	
4	58270017540	2-	2-	2-	2	1+		1.5	1.5	1.5	2	1.5	1.6	1.5	0.1	
5	58468001120	2	2	2	2	2-		2	2	2	2	1.5	1.9	2	-0.1	
6	1854001210	2+	2+	2	2+	2		2.5	2.5	2	2.5	2	2.3	2.5	-0.2	
7	58270013740	3	3	3-	3	2+		3	3	2.5	3	2.5	2.8	2.5	0.3	# 7-Details not fully elaborated. Some style.
8	57845011970	3	3+	3	3	3-		3	3.5	3	3	2.5	3	3	0	
9	58109021090	3	4-	3	3+	3		3	3.5	3	3.5	3	3.2	3.5	-0.3	
10	58459001380	3+	4-	4-	4	3		3.5	3.5	3.5	3.5	3	3.4	3.5	-0.1	
11	58468800930	4-	4	4	4	3		3.5	4	4	4	3	3.7	4	-0.3	
12	58459012340	4+	4+	4	4	3+		4.5	4.5	4	4	3.5	4.1	4.5	-0.4	
13	58270017270	5-	5	5	4+	4		4.5	5	5	4.5	4	4.6	4.5	0.1	#13 - Voice! Tone!
Discrepancy															1	
%															8%	

Comments

These papers seemed to be in order already.

Summary of Anchor Paper Review
Grade 8

Sequential Number	Paper N.	*1st Score*						*2nd Score*						*Orig.* score	*Diff* score	*Notes*
		1	2	3	4	5		1	2	3	4	5	mean			
1	58459003980	1	2	1-	1			1	1	0.5	1		0.88	1	-0.13	
2	58083002480	1	2-	1	1			1	1	1	1		1.00	1	0.00	
3	58459005660	1	1	1-	0			1	1	0.5	0		0.63	1	-0.38	
4	58083003030	1	2-	1	1+/2-			1	1	1	1.5		1.13	2	**-0.88**	
5	58109012020	2	3-	3+	3			2	2	2	3		2.25	2	0.25	
6	57854004330	3	3	3+	3			3	3	3.5	3		3.13	2	**1.13**	Did not anchor
7	58	2	3	2+	2+/3-			2	2	2.5	2.5		2.25	2	0.25	
8	58270002690	3	3	3-	3+			3	3	2.5	3.5		3.00	3	0.00	
9	58459004630	3	4	4	3+			4	4	4	3.5		3.88	4	-0.13	
10	58459003060	4-	4-	2+	3/4			3	3	2.5	3.5		2.83	3	-0.17	
11	58056008310	4	4	4	4+			4	4	4	4.5		4.13	4	0.13	
12	58088012880	3	3	3+	3+/4-			3	3	3.5	3.5		3.25	3.5	-0.25	
13	58270002510	4	4	4	4			4	4	4	4		4.00	4	0.00	
14	58459004490	4+	3-	3-	4+/5?			3	2.5	2.5	4.5		3.13	3	0.13	
Discrepancy															2	
%															14%	

Comments

Summary of Anchor Paper Review

Grade 10

Sequential Number	Paper N.	1st Round					2nd Round						Orig. score	Diff score	Notes
		1	2	3	4	5	1	2	3	4	5	mean			
1	58109014570	1	1	1	1		1	1	1	1		1	1	0	
2	57845007160	1	1	1-	1		1	1	0.5	1		0.875	1	-0.13	
3	58109014490	1	1+	1	2		1	1.5	1	2		1.375	1	0.375	
4	57854007110	1	3-	2/2+	2		1	2.5	2	2		1.875	2	-0.13	
5	58109014520	2	3+	2/2+	3-		2	3.5	2	2.5		2.5	2	**0.5**	
6	57863002400	2	3-	3-	3		2	2.5	2.5	3		2.5	2	**0.5**	
7	57854006720	2+	3+	4-	3-		2.5	3.5	3.5	2.5		3	2	**1**	
8	58459010990	2	3	4-	3-		2	3	3.5	2.5		2.75	3	-0.25	
9	58459010530	3	4	3-	3+		3	4	2.5	3.5		3.25	3	0.25	
10	57854008550	3	2	4-	3		3	2	3.5	3		2.875	3	-0.13	
11	57854007030	3	3+	3+	3		3	3.5	3.5	3		3.25	3	0.25	
12	58459008850	4	3	4+/5-	4-		4	3	4.5	3.5		3.75	4	-0.25	
13	58109015450	5	3	4+/5-	4+++		5	3	4.5	4.5		4.25	4	0.25	
14	57854006770	4	3	5	4		4	3	5	4		4	4	0	
15	57854008070	3+/4-	3	5	4-		3.5	3	5	3.5		3.75	5	**-1.25**	Not an anchor
Discrepancy															
%														4	
														27%	

Comments

I have questions about the readability / legibility of the scanned copies, the scores used "on line". Is it similar to what we see on the Score-View CD? How much of a factor is the ability to read what the kid wrote in scoring? Eligible text slows the reader's reading rate and can affect comprehension. I found some of the copies in my packet hard to read and I had to work extra hard to make sure I understood what the student was saying. How important is knowledge of the reading passage to scoring this for writing? One essay in our packet demonstrated simple control of the five paragraph essays. Writing, sentences, conventions and so forth. It wasn't great, but it was adequate. However, the essay could have been written without reading the passage. There are no specific examples from the text. Wait. I'm not sure what my question is here. In the lower-end respsonses, kids regurgitated details and made minimal or simple observations. The higher end responses included references to the text but the strength of their writing was in their observations, conclusions, insights, "particular take on the subject' rather than in the slavish retelling of details from the passage. Well, let's see what other people have to

Attachment G
Records For Re-scoring

	GR	ID	DIST	SEX	RACE	RSCAL	WRITER	WPROMP	WTEXT	RESCORE							
										1	2	3	4	5	diff	new	relia
137	3	367901	33	M	2	406	5	4	1	1	1	1	1	1	0	1	5
173	3	282450	33	M	5	374	5	4	1	1	1	1	1	1	0	1	5
176	3	366623	33	F	5	420	6	4	2	2	2	2	2	2	0	2	5
206	3	169971	33	F	2	482	7	6	1	3	3	3	3	3	+2	3	5
339	3	663722	33	F	5	464	7	5	2	3	3	3	3	3	+1	3	5
340	3	408344	33	F	5	499	6	5	1	2	3	2	2	2	+1	2	4
375	3	902913	33	M	5	442	6	5	1	2	2	2	1	2	+1	2	4
399	3	985078	33	F	5	359	4	3	1	1	1	1	1	1	0	1	5
441	3	159275	33	M	4	420	8	6	2	4	4	4	4	4	+2	4	5
511	3	549456	33	M	5	461	5	4	1	2	2	2	2	2	+1	2	5
619	3	200939	33	M	5	428	5	4	1	1	2	2	1	2	0	1	3
675	3	751398	33	F	2	436	5	4	1	2	2	2	2	2	+1	2	5
689	3	762567	33	F	5	406	7	5	2	2	2	2	2	2	0	2	5
910	3	779588	33	M	2	420	7	6	1	2	3	3	3	3	+2	3	4
927	3	106739	33	M	5	475	7	5	2	2	2	2	2	2	0	2	5
1493	3	439440	33	M	5	504	4	3	1						*		
1501	3	168315	33	F	5	532	9	6	3						*		
1509	3	506901	33	F	5	464	5	4	1	3	2	3	3	3	+2	3	4
1676	3	195108	33	M	2	431	7	5	2	2	2	2	3	2	0	2	4
1727	3	216213	33	F	2	451	7	5	2	3	3	3	3	3	+1	3	5
2006	3	230378	24	M	5	414	3	2	1						*		
2136	3	217638	23	F	2	482	9	6	3	1	1	1	1	1	-2	1	5
2234	3	75317	23	M	5	471	5	4	1	3	3	3	3	3	+2	3	5
2409	3	188615	32	F	4	326	3	2	1	2	2	2	2	2	+1	2	5
2455	3	544747	32	M	5	425	5	4	1	2	1	1	1	2	0	1	3
2477	3	78226	32	M	4	379	5	4	1		1	1	1	1	0	1	4
2526	3	49005	32	M	2	433	3	2	1	2	2	2	1	2	+1	2	3
2543	3	238821	32	M	2	362	4	3	1	1	1	1	1	2	0	1	4
2880	3	737112	32	F	2	401	5	4	1	1	1	1	1	1	0	1	5
2947	3	924873	32	F	2	379	3	2	1	1	1	1	1	1	0	1	5
2951	3	351243	32	F	5	478	7	5	2	2	2	2	2	2	0	2	5
2955	3	33818	32	F	2	448	8	6	2	2	2	2	2	2	0	2	5
3101	3	190639	32	F	5	478	7	6	1	2	2	2	2	2	+1	2	5
3190	3	150792	32	F	2	451	7	6	1	2	2	2	2	2	+1	2	5
3255	3	151192	32	M	5	396	3	2	1	1	1	1	1	1	0	1	5
3284	3	174177	32	F	5	425	3	2	1	1	1	1	1	1	0	1	5
3302	3	527531	32	M	5	349	3	2	1	1	1	1	1	1	0	1	5

3326	3	806611	32	F	2	448	7	5	2	2	2	2	2	2	0	2	5
3463	3	866307	32	M	2	401	5	4	1	1	1	1	1	1	0	1	5
3524	3	510073	32	F	4	398	5	4	1	1		1	1	1	0	1	4
3603	3	724725	32	F	5	388	5	4	1	1	1	1	1	1	0	1	5
3875	3	54689	36	F	2	417	6	5	1	1	1	1	1	1	0	1	5
3881	3	262872	36	F	2	428	3	2	1	2	2	2	2	2	+1	2	5
3899	3	324793	36	F	5	431	5	4	1	2	2	2	2	2	+1	2	5
4007	3	117443	36	M	2	359	5	4	1	1	1	1	1	1	0	1	5
4033	3	321053	36	F	5	457	6	5	1	2	2	2	2	2	+1	2	5
4093	3	100674	36	M	5	393	3	2	1	1	1	1	1	1	0	1	5
4166	3	377607	36	M	5	431	6	4	2	2	2	2	2	2	0	2	5
4209	3	780056	36	M	2	464	7	5	2	3+	3	3	3	3	+1	3	5
4274	3	445109	36	M	2	439	4	3	1	2	2	2	2	2	+1	2	5
4555	3	994327	34	M	5	433	4	3	1	1	1	1	1	1	0	1	5
4594	3	648103	34	F	5	422	4	2	2	3	2	2	2+	3	0	2	3
4848	3	882657	34	F	5	464	7	5	2	3	3	3	3	3	+1	3	5
4854	3	924279	34	F	3	464	7	5	2	3	3	3	3	3	+1	3	5
4859	3	834471	34	M	2	377	5	4	1	2	2	2	2	2	+1	2	5
4919	3	44322	34	M	5	448	6	5	1	2	2	2	2	2	+1	2	5
4951	3	795610	34	M	5	482	3	2	1	1	1	1	1	2	0	1	4
5138	3	213309	34	F	2	431	4	3	1	1	1	1	1	1	0	1	5
5468	3	104527	10	M	5	404	4	3	1	2	1	1	1	1	0	1	5
5508	3	900592	10	M	5	486	7	5	2	3	3	3	3	3	+1	3	5
5628	3	208990	10	M	5	486	7	6	1	3	3	3	3	3	+2	3	5
5634	3	344208	10	F	5	439	5	3	2	2	2	2	2	2	0	2	5
5689	3	89409	10	M	5	388	3	2	1	1	1	1	1	1	0	1	5
5763	3	494970	10	M	5	393	5	4	1	1	1	1	1	1	0	1	5
5852	3	875699	35	M	5	409	7	6	1	2	1	2	1	2	+1	2	3
6148	3	385115	18	M	5	390	4	3	1	2	1	2	1	2	+1	2	3
6162	3	335241	18	F	2	519	9	6	3	1	1	1	1	1	-2	1	5
6180	3	631982	18	M	2	417	5	4	1	3	3	3	3	3	+2	3	5
6461	3	16968	15	M	5	436	5	3	2	2	2	2		2	0	2	4
6532	3	627630	15	F	5	359	4	3	1	1	1	1		1	0	1	4
6883	3	88027	13	F	5	393	7	6	1	1	1	1	1	1	0	1	5
6895	3	662130	13	M	5	390	5	4	1	1	1	1	1	1	0	1	5
6901	3	652308	13	F	2	425	5	3	2	2	2	2	2	2	0	2	5
6984	3	675191	13	F	5	471	8	5	3	4	3	3	3	3	0	3	4
7016	3	285088	13	F	2	377	5	4	1						*		
7022	3	286314	13	F	2	420	5	3	2	2	2	2	2	3	0	2	4

7122	3	31189	29 F	5	482	6	5	1	1	2	2	2	2	+1	2	4
7124	3	135902	29 M	5	442	6	4	2	2	2	2		2	0	2	4
7146	3	330812	29 F	5	422	5	4	1	2	1	1	1	2	0	1	3
7224	3	382099	29 F	5	338	3	2	1	1	1	1	1	1	0	1	5
7295	3	968646	29 M	5	467	5	4	1	2	3	3	3	3	+2	3	4
7437	3	17259	31 M	5	442	7	5	2						*		
7481	3	506740	31 F	2	471	9	6	3	2	2	2	2	2	-1	2	5
7553	3	185970	31 M	5	490	7	6	1	1	1	1	1	1	0	1	5
7672	3	885879	31 M	2	422	4	3	1	1	1	1	1	1	0	1	5
7673	3	14769	31 F	2	388	3	2	1	2	3	3	2	3	+2	3	3
7686	3	107885	31 F	5	454	7	5	2	1	1	2	1	2	0	2	3
7722	3	177359	31 M	5	454	7	6	1	2	3	3	3	3	+2	3	4
7758	3	329557	31 F	3	464	7	5	2	1	1	1	1	1	-1	1	5
7786	3	915274	31 F	5	442	7	6	1	2	1	1	1	2	0	1	3
7827	3	530891	31 M	2	436	9	6	3		2	3		2	-1	2	2
7859	3	33811	31 M	5	525	6	4	2		2			3	0	2	1
7955	3	916639	31 M	5	454	4	3	1	2	2	2	2	2	+1	2	5
7961	3	648639	31 M	5	514	7	5	2		3	3	4	4	+1	3	2
7982	3	472929	31 M	5	433	6	5	1		1	1	1	1	0	1	4
7990	3	623723	31 M	2	404	6	5	1		1	1	1	1	0	1	4
8062	3	357918	31 F	5	439	8	6	2		2	2	2	2	0	2	4
8107	3	95736	31 M	5	422	5	4	1		1	1	1	1	0	1	4
8115	3	744756	31 F	2	409	8	6	2		2	2	2	2	0	2	4
8385	3	973827	17 F	5	464	7	5	2			2			0	2	1

Numb of ca read: 10 0 N mber of cas es list ed: 100

	GR	ID	DIST	SEX	RACE	RSCAL	WRITERA	WPROMP	WTEXT	RESCORE							
										1	2	3	4	5	diff	new	relia
8779	5	75336	33	F	5	502	7	5	2	1	1	1	1	1	-1	1	5
8792	5	827470	33	F	5	419	5	4	1	1	1	1	1	1	0	1	5
9156	5	267652	33	F	2	411	5	4	1	1	1	1	1	1	0	1	5
9179	5	75485	33	F	2	467	8	6	2	2	2	2	3	2	0	2	4
9488	5	499368	33	F	2	498	8	6	2	2	2	1	1	2-	0	2	3
9504	5	649423	33	M	5	513	7	4	3	3	3	3	4	3	0	3	4
9528	5	404329	33	M	5	470	6	5	1	2	2	2	2	2	1	2	5
9619	5	373309	33	M	5	473	6	5	1	2	2	2	1	2	1	2	4
9634	5	861363	33	M	2	436	7	6	1	2.5	3	2	3	3-	2	3	3
9643	5	942274	33	M	5	434	5	4	1	1	1	2	1		0	1	3
9652	5	206655	33	M	5	465	7	5	2	2	2	2	3	2	0	2	4
9771	5	995356	33	M	2	473	10	7	3	0	0	0		0	0	3	4
9822	5	288141	33	F	5	495	6	4	2	2	2	2	1	2	0	2	4
9922	5	741739	33	F	4	444	8	6	2	1+		1	1		0	1	3
9971	5	290252	33	F	5	495	5	4	1	2		2	1	2	1	2	3
10047	5	708832	33	M	2	401	4	3	1	1		1	1	1	0	1	4
10096	5	416506	33	F	5	488	5	4	1	1		0	0		-1	0	2
10126	5	340542	33	F	2	444	8	6	2	2		2	3	2	0	2	3
10162	5	684925	33	F	5	556	11	7	4	3		3	3	3	-1	3	4
10393	5	491280	24	F	5	485	7	6	1	2		1	1		0	1	2
10421	5	542476	24	M	5	491	7	5	2	2		2	1	2+	0	2	3
10422	5	866434	24	F	5	518	5	4	1	2		2	2		1	2	3
10467	5	925188	24	F	5	505	10	7	3	3		3	4		0	3	2
10483	5	633385	24	M	3	441	6	5	1	2-		1	1		0	1	2
10506	5	787171	24	M	5	479	8	6	2	2		2	2		0	2	3
10671	5	193079	23	M	2	470	6	5	1	2		2	2		1	2	3
10758	5	174997	23	M	5	424	5	4	1	1	1	1		1	0	1	4
10814	5	193758	23	M	5	505	8	6	2	2	2	1		2	0	2	3
11025	5	118792	32	M	5	549	7	5	2	3	3+	1		3	1	3	3
11118	5	379913	32	F	2	429	4	3	1	2-	1	3		1	0	1	2
11239	5	63303	32	M	5	476	5	4	1	2-	1	1		1	0	1	3
11314	5	204606	32	M	2	522	7	6	1	3	4 3	3		3-	2	3	4
11509	5	406330	32	M	5	429	5	4	1	1	1	1		1	0	1	4
11562	5	201850	32	M	5	467	8	6	2	2	2	1		2+	0	2	3
11583	5	257101	32	F	2	403	6	5	1	1	2	1		2	0	1	2
11597	5	739867	32	F	5	467	5	4	1	2	1	1		2-	0	1	2
11716	5	210501	32	M	5	502	8	6	2	2	3	3		3	1	3	3

11788	5	177380	32 F	5	527	8	5	3	3	4-	3		3	0	3	3
11876	5	55245	32 M	2	406	5	4	1	1	1	1		1	0	1	4
11923	5	259036	32 F	4	498	8	6	2	3	2+	2		3	0	2	2
11999	5	755348	32 M	5	482	3	2	1	1+	1	1		1+	0	1	4
12119	5	805167	36 F	5	491	7	6	1	2	2	1		2	1	2	3
12226	5	795689	36 M	5	414	3	2	1	1	1	1		1	0	1	4
12348	5	862536	36 M	5	462	3	2	1	1	1	1		1	0	1	4
12417	5	700730	36 F	5	522	8	6	2	2+	3	2		3	0	2	2
12594	5	828385	36 F	5	518	8	6	2	3+	4	3		4-	1	3	2
12726	5	43250	34 F	2	495	7	5	2	3+	3	3		3	1	3	4
12824	5	142480	34 F	2	389	6	5	1	1	1	1		1	0	1	4
12845	5	797280	34 M	2	446	5	4	1	2	2	1		2-	1	2	3
12952	5	572284	34 F	2	436	8	6	2	2	2	1		2	0	2	3
12966	5	10637	34 M	2	462	8	5	3	2	2	1		2	-1	2	3
13027	5	888974	34 M	5	502	6	4	2	2	2	2		2	0	2	4
13113	5	231993	34 F	4	414	6	5	1	1	1	1		1	0	1	4
13320	5	176890	34 F	5	470	8	6	2	2		2	2		0	2	4
13358	5	384747	34 M	5	462	4	3	1	2		2	2		1	2	3
13444	5	560812	34 F	2	485	5	3	2	2		2	1		0	2	2
13450	5	196784	34 F	5	479	7	5	2	2		3	2		0	2	2
13723	5	360070	10 M	3	465	5	4	1	1+		1	1		0	1	3
13825	5	974179	10 M	2	409	3	2	1	1		1	1		0	1	3
13837	5	487971	10 F	5	518	8	6	2	3		3	2		1	3	2
13881	5	654538	10 M	5	488	6	5	1	1		1	1		0	1	3
13950	5	471218	10 M	2	457	5	4	1	1		1	1		0	1	3
14154	5	177031	35 M	2	462	5	4	1	2		2	2		1	2	3
14201	5	339384	35 F	4	473	7	5	2						*		
14262	5	206058	18 F	5	502	9	6	3	3		3-	3		0	3	3
14338	5	394600	18 M	5	427	7	5	2	1		1	1		-1	1	3
14551	5	317717	16 F	5	451	4	3	1	2	1+			2-	1	2	2
14713	5	825086	15 F	5	459	6	4	2	2	2			3-	0	2	2
14732	5	115215	15 M	5	527	9	7	2	3	3			4	1	3	2
14737	5	220091	15 M	5	439	6	4	2	2	2			2	0	2	3
14843	5	95574	15 M	5	488	7	5	2	1	2			2-	1	2	2
14896	5	230189	15 M	5	476	5	4	1	2	1			2-	1	2	2
14904	5	768636	15 M	2	467	8	6	2	2	3			2	0	2	2
14954	5	120990	13 M	2	377	3	2	1	1					0	1	1
14985	5	501914	13 M	2	446	6	5	1						*		
15119	5	501759	13 M	5	446	3	2	1		1			1	0	1	2

15235	5	375826	13 F	5	498	7	5	2	2	2			1+ 2-	0	2	3
15241	5	895376	13 F	5	513	7	5	2	2	2			3-	0	2	2
15331	5	301588	13 M	2	476	4	3	1	1	1			1	0	1	3
15357	5	350104	13 M	5	527	10	7	3	3	3	3	2	3	0	3	4
15455	5	57488	29 M	3	470	7	5	2	2	2	1	1	2	0	2	3
15626	5	125143	29 M	5	465	5	4	1	1	1	1	1	1	0	1	5
15682	5	78300	29 F	5	454	6	5	1	1	1+	0	0	1	0	1	3
15784	5	205098	29 F	5	467	5	4	1	2	2	1	1	1	0	1	3
15958	5	241926	31 F	5	454	6	4	2	2	2	2	1	2-	0	2	4
15983	5	706099	31 M	2	441	6	5	1	1	1	1	1	1	0	1	5
16002	5	497317	31 M	5	457	5	4	1	2	2	2	1	2	1	2	4
16004	5	964751	31 M	4	495	8	6	2	4	4	4	4	4	2	4	5
16085	5	486233	31 F	5	556	9	6	3	4	3	4	4	4	1	4	4
16095	5	167048	31 F	5	543	10	7	3	5	5	4	4	5	2	5	3
16132	5	735548	31 M	5	451	7	5	2	2+	2	2	2	3-	0	2	4
16155	5	948607	31 F	2	444	8	6	2	2+	3	2		2	0	2	2
16265	5	147686	31 F	5	439	6	4	2	3	3	3		3	1	3	4
16339	5	343527	31 F	2	488	9	6	3	3		3		4-	0	3	2
16395	5	498619	31 F	5	543	10	7	3	3		3-		3	0	3	3
16508	5	505905	31 F	2	482	6	5	1	1		1		1+	0	1	3
16523	5	222319	31 M	5	549	9	6	3	2+		2		3	1	2	2
16630	5	595315	31 F	5	543	8	6	2	3		2		3-	1	3	2
16940	5	772364	17 M	5	532	8	6	2	2		2		3-	0	2	2
17017	5	680130	17 M	5	479	5	4	1	2		2		3+	1	2	2

Number of ca read: 10 0 Nu mber of cas es list ed: 100

	GR	ID	DIST	SEX	RACE	RSCA	WRITERA	WPROMP	WTEXT	RESCORE							
										1	2	3	4	5	diff	new	relia
17077	8	937445	33	M	5	532	8	7	1	2	1 2	2			1	2	3
17226	8	214817	33	F	5	532	7	5	2	2	1 2	3			0	2	2
17346	8	884438	33	F	5	556	8	5	3	3	4	3			0	3	2
17582	8	924867	33	M	5	513	8	6	2						*		
17844	8	673683	33	M	4	490	8	6	2	2	2 3	2			0	2	3
17996	8	329707	33	F	4	552	8	5	3	3	4	3			0	3	2
18009	8	940992	33	M	5	502	7	6	1	1	1	1			0	1	3
18056	8	924818	33	M	5	532	7	5	2	3	3	4			1	3	2
18060	8	493567	33	F	5	524	10	7	3	3	3 4	4			0	3	2
18067	8	192943	33	M	2	492	6	5	1						*		
18211	8	916864	33	F	4	521	6	5	1	2	1	2			1	2	2
18377	8	307335	33	F	5	535	8	6	2	2	1 2	2			0	2	3
18416	8	752786	33	F	5	513	7	6	1	2	1 2	2			1	2	3
18494	8	690583	33	F	5	487	8	6	2	2	1	1		1	-1	1	3
18725	8	593665	24	M	2	510	7	6	1	3	2	2		2	1	2	3
18827	8	306480	24	F	5	515	8	6	2	2	3	3		2	0	2	2
19030	8	860137	23	M	5	548	7	5	2	2	3	3			1	3	2
19036	8	231801	23	F	2	480	8	6	2	2	3	2			0	2	2
19098	8	649734	23	M	2	532	9	7	2	2	3	2			0	2	2
19179	8	193408	32	F	2	468	8	6	2	2	2+3-	3			0	2	2
19182	8	431133	32	M	5	444	7	6	1						*		
19239	8	639143	32	F	2	542	5	4	1	1	1	1			0	1	3
19288	8	758424	32	M	5	450	6	5	1	1	1	0			0	1	2
19863	8	312403	32	F	5	500	5	4	1		0				0	1	1
20147	8	593838	32	F	5	497	6	5	1	1	1	1			0	1	3
20148	8	830675	32	M	5	480	7	6	1	2	2	2			1	2	3
20153	8	699218	32	M	5	524	8	6	2	2	2	2			0	2	3
20162	8	403732	32	F	5	505	9	7	2		2	2			0	2	2
20380	8	740581	32	M	5	529	6	4	2		1 2	2			0	2	2
20528	8	942121	36	M	5	532	7	5	2	2	2	2			0	2	3
20553	8	677223	36	M	4	468	6	4	2	2	3-	3			1	3	2
20581	8	672452	36	M	5	552	8	6	2	4	4	4			2	4	3
20627	8	456895	36	F	5	545	10	7	3	4	4	4			1	4	3
20703	8	44452	36	F	2	577	9	6	3		3	3			0	3	2
20757	8	721301	36	M	5	458	6	4	2		2	0			0	2	1
20759	8	771125	36	M	5	545	8	6	2		3	3			1	3	2
20794	8	686203	36	F	5	510	8	6	2		2	2			0	2	2

21010	8	234133	36 F	5	542	9	7	2		2	2			0	2	2
21066	8	596917	97 M	5	500	7	6	1		1+ 2	1			0	1	2
21119	8	269698	34 M	5	515	8	7	1		1 2	2			1	2	2
21130	8	951273	34 F	2	483	5	4	1	2 3	3	3	3	3	2	3	5
21132	8	104065	34 F	2	441	4	3	1	2	2	2	3	2	1	2	4
21212	8	108503	34 M	5	470	5	4	1	1	1	1	1	1	0	1	5
21215	8	261052	34 M	5	502	8	6	2	1	1	1	1	1	-1	1	5
21285	8	987757	34 M	5	615	9	6	3	2	2	1	2	2	-1	2	4
21352	8	50671	34 F	5	526	8	5	3	4	4+	4	4	4	1	4	5
21426	8	350461	34 M	5	535	7	5	2	2	2	0	2	0 2	0	2	4
21614	8	110646	34 F	5	542	8	6	2	2	2	2		3	0	2	3
21625	8	676398	34 F	2	538	10	7	3	3	3	4		3	0	3	3
21760	8	473838	34 F	5	521	7	6	1	3	3	3		3	2	3	4
21780	8	784723	34 F	5	529	7	6	1	2	2			2	1	2	3
21823	8	416918	34 M	2	548	8	6	2	2	3			2	0	2	2
21846	8	143951	34 F	5	444	5	4	1	1	0			0 1	0	1	2
22130	8	151451	10 F	5	568	8	6	2	2			3		0	2	1
22461	8	276423	10 F	2	510	9	6	3						*		
22476	8	279257	10 M	5	473	7	5	2						*		
22488	8	183440	10 F	5	529	8	7	1						*		
22598	8	855969	35 F	2	485	7	6	1	1			2	1	0	1	2
22604	8	922297	35 F	4	535	6	4	2	2			3	3	1	3	2
22620	8	815644	35 F	5	502	6	5	1	2			2	2	1	2	3
22694	8	175531	71 M	5	473	4	3	1	2			2	2	1	2	3
22808	8	317809	18 F	5	485	6	4	2	1			1	1	-1	1	3
22855	8	63650	18 M	5	518	5	4	1	3			3	3	2	3	3
22869	8	226118	18 M	5	545	8	6	2	2			2	2	0	2	3
23054	8	69546	18 F	5	572	10	7	3	4			4	4	1	4	3
23197	8	773717	16 F	5	526	7	5	2	2			3	3	1	3	2
23369	8	480151	15 M	3	548	9	7	2		2+ 3-	3	3	3	1	3	3
23442	8	733068	15 F	5	510	8	6	2		2+ 3-	2	3	2+	0	2	3
23481	8	105320	15 F	2	500	8	6	2		2	1	2	2+	0	2	3
23596	8	536530	37 F	5	500	8	6	2		2	1	2	2	0	2	3
23640	8	512019	37 M	5	538	7	6	1		2 0	1	2	2	1	2	3
23771	8	393311	13 M	2	524	7	5	2		2	1	2	2	0	2	3
23782	8	336343	13 F	2	526	8	6	2		3	2	3	3	1	3	3
23854	8	450678	13 M	2	532	5	4	1		1	2	2	2	1	2	3
23965	8	790435	13 M	5	450	4	3	1		2-	2	2	2	1	2	4
24173	8	557401	29 F	5	577	9	7	2		2+	3	3	3	1	3	3

24314	8	887415	29 M	5	532	8	6	2		2	2	2	2	0	2	4
24422	8	117058	29 M	5	505	6	5	1		2	1	2	2	1	2	3
24487	8	129208	29 M	5	502	6	5	1		2	1 2	3	3	2	3	2
24557	8	36489	31 F	2	444	7	5	2		0 3		2	2	0	2	2
24578	8	883430	31 M	5	505	8	6	2	2			2	2	0	2	2
24611	8	144665	31 M	5	564	10	7	3				4	4	1	4	2
24624	8	264936	31 M	5	529	6	4	2	2			3	3	1	3	2
24662	8	215581	31 M	5	577	9	6	3	3			3	3	0	3	2
24706	8	952232	31 F	5	524	9	6	3	2				4 5	1	4	1
24723	8	852922	31 F	2	559	10	7	3	3			4	3 4	0	3	2
24935	8	468679	31 F	3	548	7	4	3	2			2	2	-1	2	3
25076	8	456464	31 F	5	480	5	4	1	2			4	3 4	3	4	2
25317	8	731942	31 M	5	593	7	5	2	2				3 4	1	3	1
25338	8	748164	31 M	5	577	8	6	2	2			3	3	1	3	2
25463	8	296624	31 F	5	532	9	7	2	2			3	3	1	3	2
25573	8	577938	17 M	5	552	7	6	1	3			3	2 3	2	3	3
25582	8	24983	17 M	5	524	8	6	2	3			4	3+	1	3	2
25589	8	475989	17 M	5	529	7	6	1	1		1	2	2	0	1	2
25640	8	61135	17 M	5	505	7	6	1	1		2	2	2	1	2	3
25655	8	797223	17 F	5	485	5	4	1	1		1	2	2	0	1	2
25707	8	931228	17 M	5	545	8	6	2	2		4	5	4 5	2	4	2
25781	8	942209	17 F	5	513	7	5	2	2		4	2	2	0	2	3
25805	8	234851	17 F	5	552	9	7	2	2		3	3	3	-1	3	3
25871	8	437136	17 F	2	524	8	6	2	4		4	3	3	1	3	2

Number of ca read: 10 0 N mber of cas es list ed: 100

	GR	ID	DIST	SEX	RACE	RSCALE	WRITERA	WPROMP	WTEXT	RESCORE						DIFF	new scor	relia
										1	2	3	4	5	6			
26015	10	419991	38	F	2	527	8	6	2			2		2	2	0	2	3
26039	10	421647	38	F	2	509	6	5	1			2		2	2	1	2	3
26105	10	652857	38	F	5	512	6	4	2			2		2	2	0	2	3
26134	10	425833	38	M	5	512	9	7	2			3		3	3	1	3	3
26210	10	796376	38	F	5	537	8	6	2			3		3	3	1	3	3
26227	10	971450	38	F	5	544	7	5	2			2		2	2	0	2	3
26344	10	964948	38	F	5	479	7	5	2			2		2	2	0	2	3
26538	10	211913	38	F	5	476	7	5	2			2		2	2	0	2	3
26567	10	500437	38	F	5	524	8	6	2			3		3	3	1	3	3
26613	10	350650	38	M	5	503	8	6	2			2		2	2	0	2	3
26652	10	908659	38	F	2	509	10	7	3					3	4	0	3	1
26807	10	425261	33	F	5	563	9	6	3					3	4	0	3	1
26831	10	841297	33	M	5	540	10	7	3					3	3	0	3	2
26873	10	773780	33	F	5	537	7	4	3	3	3	3		3	3	0	3	5
26968	10	892106	33	F	5	600	11	7	4	3	4	4		4	4	0	4	4
26989	10	779708	33	M	5	537	8	6	2	2	2	2		2	2	0	2	5
27105	10	18948	33	F	5	521	9	6	3	2	2	2		3	2	-1	2	4
27108	10	19499	33	F	5	559	9	6	3	3	3	3		3	3	0	3	5
27149	10	923401	33	M	5	607	8	6	2	3	4	4		3	4	2	4	3
27212	10	877800	33	M	2	487	5	4	1	1	1	1		2	1	0	1	4
27329	10	912641	33	F	5	506	8	6	2	1	1	1		1	1	-1	1	5
27348	10	629522	33	F	5	481	5	4	1	1	1	1		1	1	0	1	5
27434	10	674010	33	M	2	481	5	4	1	1	1	1		2	1	0	1	4
27435	10	892	33	F	5	537	5	4	1	2	2	2		2	2	1	2	5
27495	10	228356	33	F	5	551	6	4	2	2	2	2			2	0	2	4
27521	10	519860	33	M	5	479	6	5	1	0	0	0		1	0	0	1	4
27609	10	232470	33	M	5	518	9	7	2	3	3	3		2	3	1	3	4
27662	10	216344	33	F	5	503	4	2	2	2	3	3		3	3	1	3	4
27702	10	446515	33	M	5	428	5	4	1	0	0	0		0	0	0	1	5
27823	10	364707	33	F	5	415	3	2	1	1	0	0		1	0	-1	0	3
27840	10	77470	33	F	3	521	7	5	2	2	3	3		3	3	1	3	4
27914	10	663861	40	M	5	540	8	6	2	3	3	3		2	3	1	3	4
28017	10	263219	40	M	5	495	7	6	1	2	2	2			2	1	2	4
28159	10	426000	40	M	2	449	7	6	1	1	1	1		1	1	0	1	5
28357	10	782281	24	M	5	530	5	4	1	1	2	2		2	2	1	2	4
28599	10	491714	23	M	5	551	8	6	2							*		
28753	10	796778	32	M	5	473	5	4	1			0		0	0	-1	0	3

	GR	ID	DIST	SEX	RACE	RSCALE	WRITERA	WPROMP	WTEXT			RESCORE				DIFF	new scor	relia
										1	2	3	4	5	6			
28782	10	559814	32	F	5	537	6	5	1	2		2		2	2	1	2	4
28806	10	960590	32	M	4	441	4	3	1	1		1		1	1	0	1	4
28819	10	881461	32	M	5	600	8	6	2	3		4		4	4	2	4	3
28838	10	352215	32	M	5	500	5	4	1	2	1			2	2	1	2	3
29096	10	34839	32	M	5	455	5	4	1	0	0			0	0	-1	0	4
29129	10	953257	32	F	5	572	11	8	3	2	2			2	2	-1	2	4
29346	10	595	32	F	5	518	9	6	3	2	2			1	2	-1	2	3
29618	10	661766	39	M	5	495	6	5	1	1	1			1	2	0	1	3
29627	10	376334	39	M	4	506	8	6	2	2	2			2	3	0	2	3
29653	10	982616	39	M	5	455	7	6	1	2	2			2	2	1	2	4
29800	10	389383	36	F	5	582	10	6	4	4	3			4	4	0	4	3
29810	10	152017	36	M	5	503	8	5	3	2	3			3	3	0	3	3
29854	10	763885	39	M	2	498	5	4	1							*		
29965	10	760851	36	F	1	582	10	7	3	4	4			4	4	1	4	4
29995	10	870781	36	F	5	487	6	4	2	2	2			2	2	0	2	4
30091	10	913782	36	F	5	572	9	8	1	3	3			2	3	2	3	3
30112	10	580310	36	F	5	537	8	6	2	2	1					0	2	2
30186	10	898181	36	F	2	498	8	6	2	2	1					0	2	2
30502	10	402309	34	F	5	567	9	6	3	2	2					-1	2	3
30582	10	878046	34	M	2	492	8	6	2	1	1		1			-1	1	3
30834	10	925911	34	F	5	521	8	6	2	2	2					0	2	2
30853	10	103251	34	F	2	537	9	6	3	3	3					0	3	2
30904	10	187357	10	F	5	489	7	6	1							*		
30929	10	990088	10	M	2	500	8	6	2	3	3					1	3	2
30991	10	294931	10	F	4	521	8	6	2	1	1					-1	1	2
31134	10	615078	10	M	2	463	3	2	1	1	1					0	1	2
31191	10	610598	10	M	5	515	6	4	2							*		
31193	10	660459	10	M	5	600	9	8	1	1	1					0	1	2
31326	10	206564	35	M	5	563	9	6	3	2	2					-1	2	2
31358	10	293609	35	F	5	518	6	4	2	3	3			3	3	1	3	4
31617	10	376403	18	F	4	481	5	4	1	3	3		3	3	3	2	3	5
31652	10	23777	16	M	5	438	6	5	1	3	3			3	4	2	3	3
31721	10	92297	16	F	2	457	5	4	1	2	1			2	2	0	1	3
31756	10	934124	16	F	2	487	7	5	2	2	1			2	2	0	2	3
31869	10	107382	15	M	5	527	5	4	1	1	1			2	1	0	1	3
31881	10	135887	15	F	5	487	6	5	1	1	1			1		0	1	3
31882	10	761319	15	M	5	479	7	5	2	2	1			2	2	-1	1	3

	GR	ID	DIST	SEX	RACE	RSCALE	WRITERA	WPROMP	WTEXT	RESCORE						DIFF	new scor	relia
										1	2	3	4	5	6			
31961	10	109773	37	M	5	489	7	5	2	2	1			2	2	0	2	3
31993	10	673931	37	M	5	509	6	5	1	2	1		2	3	2	1	2	3
32096	10	709947	70	M	5	555	8	6	2	3	3			3	4	1	3	3
32115	10	298406	70	M	3	582	12	9	3	3	2		2	2	3	-1	2	3
32117	10	4537	70	F	2	521	9	6	3							*		
32133	10	329583	70	F	2	527	9	6	3	3	2			3	3	0	3	3
32141	10	500167	70	M	5	509	9	7	2							*		
32238	10	106376	70	M	5	537	8	6	2					3		1	3	1
32280	10	322119	13	M	5	452	7	5	2					2		0	2	1
32295	10	295461	13	F	5	527	6	4	2					2		0	2	1
32412	10	153750	13	M	5	555	11	8	3					3		0	3	1
32464	10	840910	13	M	5	492	6	4	2							*		
32814	10	616254	29	F	5	547	7	5	2					3		1	3	1
32825	10	639304	29	F	2	463	4	3	1					1		0	1	1
32861	10	981248	29	F	5	563	9	6	3					3		0	3	1
33062	10	600358	31	M	5	524	5	4	1					2		1	2	1
33188	10	214833	31	F	5	551	8	6	2					3		1	3	1
33334	10	409446	31	M	5	484	7	5	2					2		0	2	1
33383	10	246632	31	M	5	495	9	6	3					2		-1	2	1
33388	10	148398	31	F	5	506	11	7	4					3		-1	3	1
33446	10	110457	31	F	2	498	7	5	2					3		1	3	1
33500	10	597104	31	F	2	518	8	6	2					3		1	3	1
33552	10	386053	31	F	2	572	10	7	3					4		1	4	1
33817	10	631370	17	F	5	498	5	4	1					3		2	3	1
33863	10	120670	17	F	5	527	7	5	2					3		1	3	1
33896	10	296373	17	M	2	551	8	6	2					3		1	3	1

Number f cas read: 10 0 N mber of cas es list ed: 100

Attachment H
Writing Task Evaluation

Writing Task Evaluation								
17-Aug-00								
Grade 3								
	1999				2000			
Reviewer	Carter	Whitman	Trainer	Shockley	Carter	Whitman	Trainer	Shockley
							√	√
Very Hard								
Hard					√			
Not Hard	√	√	√	√		√		
Easy								
Very Easy								
Performance								
5	0.02	0.03	0.05	0.08	0.00	0.02	0.01	0.00
4	0.30	0.15	0.05	0.15	0.15	0.13	0.05	0.10
3	0.25	0.50	0.40	0.32	0.30	0.25	0.30	0.23
2	0.23	0.22	0.30	0.30	0.25	0.40	0.44	0.42
1	0.20	0.10	0.20	0.15	0.30	0.20	0.20	25.00

Writing Task Evaluation										
17-Aug-00										
Grade 5										
	1999					2000				
Reviewer	Sheets	Poehlmann	Hodgkins	Currie	Lam	Sheets	Poehlmann	Hodgkins	Currie	Lam
Very Hard										
Hard						√	√	√	√	√
Not Hard	√	√	√	√	√					
Easy										
Very Easy										
Performance										
5	0	0.05	0.02	0.00	0.05	0.00	0.01	0.02	0.00	0.02
4	0.10	0.05	0.15	0.10	0.20	0.05	0.05	0.10	0.05	0.18
3	0.25	0.50	0.38	0.40	0.35	0.20	0.30	0.30	0.20	0.20
2	0.45	0.30	0.30	0.45	0.30	0.35	0.32	0.38	0.50	0.40
1	0.20	0.10	0.15	0.05	0.10	0.40	0.32	0.20	0.25	0.20

Writing Task Evaluation										
17-Aug-00										
Grade 8										
	1999					2000				
Reviewer	Speicher	Szegda	Ash	Poorman	Angstadt	Speicher	Szegda	Ash	Poorman	Angstadt
Very Hard										
Hard		√		√			√	√	√	√
Not Hard					√	√				
Easy			√							
Very Easy										
Performance										
5		0.05	0.12	0.03		0.01	0.04	0.01	0.03	0.02
4		0.15	0.30	0.17		0.10	0.16	0.12	0.07	0.07
3		0.40	0.30	0.40		0.60	0.50	0.40	0.40	0.20
2		0.20	0.18	0.30		0.19	0.20	0.20	0.40	0.21
1		0.20	0.10	0.10		0.10	0.10	0.11	0.10	0.05
								.06 off topic		

Writing Task Evaluation										
17-Aug-00										
Grade 10										
	1999					2000				
Reviewer	Boyd	Grudzina	Laverick	Smith	Wilson	Boyd	Grudzina	Laverick	Smith	Wilson
Very Hard								√	√	
Hard	√					√				√
Not Hard		√	√	√	√					
Easy							√			
Very Easy										
Performance										
5	0.03	0.05	0.05	0.00	0.00	0.02	0.07	0.00	0.00	0.00
4	0.16	0.20	0.15	0.20	0.10	0.15	0.25	0.05	0.01	0.05
3	0.35	0.60	0.60	0.30	0.30	0.33	0.50	0.40	0.34	0.25
2	0.25	0.20	0.10	0.30	0.40	0.30	0.10	0.30	0.30	0.40
1	0.21	0.05	0.10	0.20	0.20	0.20	0.08	0.25	0.35	0.30

Comments:

- Confusing for students. It arises that tigers don't live in forests with foxes. The wording for the 1999 prompt explained it better than in the 2000 prompt did for 3rd graders to understand. The 1999 story was more interesting. Children can relate to those pets.

- The question doesn't give enough stem to make students relate to the nature of the folk tale.

- I think it would be very hard for a child to achieve a 4 or 5 on this task because if they respond to the prompt and the reading, it's basically a retelling. There is little opportunity to demonstrate any writing "flair" they may have which can push a paper to a 4 or 5.

Comments

- Felt the students needed an organizer.
- The 2000 test is harder.

Comments

- Students have some difficulty with "heritage". The question implies that the author and narrator are the same person and they are not.

- I'm not sure that this means anything to me, really. I felt like I was answering without a lot to go on. My low estimate is based on confusion based on the prompt. It assumes/implies that the author and the narrator are the same when we teach them. That is not the case. It forces the reader/writer to shift between viewpoint and voice, and that kind of stuff that can confuse issues.

Comments

- I think that the Cinderella question was not very engaging to the male population.

- This is most likely an overly biased or subjective evaluation. I'm not sure I have a good idea of an "average" student.

- I feel that the 2000 prompt uses language ("broad appeal") that is confusing to the students. Many did not seem to understand that broad appeal meant "popular" which influenced their answer.

- The word "choice" in the question (2000 essay) make it confusing for the general student population. I would re-word "broad appeal". The 1999 question had more familiar wording and topic choice. The selection was more interesting to students. The 2000 selection was more literary, and students in grade 10 received less instruction in literature.

- Perhaps the phrase "broad appeal" confuses some kids who cannot understand what it meant.

- The format of the ? {Last on page, no bold in directions, no drafting} leads itself to the difficulty level.

- 2000 – I think that the Cinderella question was not very engaging to the male population.

Attachment I
Writing Study Group Discussion

Writing Study Group Discussion
Grade 3

Text-Based Writing Development

Look at choice of text. There is a big difference between 1999 and 2000. It is much easier for kids to connect to a person rather than a monkey. Prefer to see realistic stories used when kids have to be a Text-based writing. Authentic audience.

The question in 2000 did not say use details and information for the text. The 1999 did not say that verb tense (present) may have confused them. The question doesn't give enough stem to make students relate to the nature of the folk tale.

Basically a retelling, so it is difficult for a child to achieve a 4 or 5. The question asked what do you see, what do you think and lots of kids responded, I see, I see, I see or I think. Confusion for students arises that tigers don't live in forests with foxes.

Text-Based Writing Score

We found a big discrepancy between Harcourt's and ours. They seem to have a bigger range for a 1. Check the discrepancies between our "3's" and their same pieces which were scored "1".

We also had a problem with the invalids "the big rock" scored a 1, should it be a 0?

Teachers should be given a breakdown of "the score". This code can be done by codes. Not enough details for the story, etc...

Text-based scoring rubric for writing be developed. How much emphasis is incorporated in the grading for using the text?

Test Administration

Certainly not 2 Text-based writing's on one day. Not authentic. Use a realistic story, not a fantasy or folk tale. Word the prompt more carefully – scaffold. Set up like the prompt is to emphasize the importance.

Test Development

Writing Study Group Discussion
Grade 3

DSTP Manual 2000. Page 17

The Directions include both questions 57/67 to take a 10 minute rest. Included "Dear Mrs. Cabot"

Writing Study Group Discussion
Grade 5

Text-Based Writing Development

Text-based writing added to get a more valid score for writing. 1 page versus 2 pages formatting. Change in short answer.

Navaho selection is not as engaging. Don't have background experience. Should be able to get enough from text. (But familiarity helps connect to text.)

Do better with nature. Students connect. Both informative pieces (Science and Social Studies).

Text-Based Writing Score

Text-based writing used writing rubric. Need anchor papers. Is our scoring as good without anchors as theirs is?

Test Administration

Instructions: Developmentally need to be aware of problems, of directions, understand graded for reading and writing. Amount of instruction needed to listen and to understand. Not doing 2 text-based writing's on the same day. Consider doing text-based writing passage first of a set. Prewriting area.

Classroom Instruction

Develop text-based responses to go home with class take home books. Content area do also. Model text-base responses show, have to make connections and pull out details.

Make connections with characters in the story.

New reading people to look at district curriculum to find samples of things to work on.

Grade levels / districts developing questioning activities to improve text-based writing. Sample papers for text-based writing.

Writing Study Group Discussion
Grade 5

Test Development

DSTP Test 2000 – Student Book 22 – Response booklet 12-13 Use 2 full pages. 1999 only 1 page.

Grade 5 directions read much better than grade 3. Especially need to look at breaks in grade 3.

Writing Study Group Discussion
Grade 8

Text-Based Writing Development

350461 – Issues: Is it on topic? Should it be scored?
143951 – Topic issues. Is this scorable?
11306480 – Tuff to decide between 2+/3.
193408 – 2/3---
758424 – Off topic?
721301 – Is this off topic?

What is the purpose of text-based writing, is it to test for ability to write an audience purpose? "It didn't count".

1. When is a piece off topic? {and/or invalid}. Is a summary that doesn't relate back to the question scorable? When does 1999 tell us?
2. How did media information one week prior to test impact scores?
3. Stand-alone started in 90 text-based writing… 98 not as much experience.
4. Author/narrator: Confusion or attention to author craft?
5. Did kids understand the concept of heritage?

Text-Based Writing Score

Does a narrative text lead itself to a persuasive prompt?

Writing Study Group Discussion
Grade 10

Text-Based Writing Development

1999 Text-based writing prompt is more opinion based and more authentic.
2000 prompt has less room for creativity. (Harcourt wrote it).

More like a SOAP prompt. (Subject O and purpose)

Use "user accessible" language

Doug says there should be more bone oriented Civil War passages.

Text-Based Writing Score

Relation to off topic. Should the use of the word "Cinderella" earn credit? (Re-examine the "invalids"). Insufficient information.

0-1 distinction (May be something that changes per prompt).

Test Administration

Prompt set up alone. Not the last to do in the day. Bold it . Emphasize it's importance. Separate page. Bolded for emphasize. Physically make it look like the stand alone prompt.

For Special Ed. – In unclear directions, can there be another version (allowable modification) of the directions?

Classroom Instruction

Research and Instruction – When appropriate to paraphrase/quote from a source.

T – have gone away from literary analysis due to emphasis on stand-alone prompts.

Variety of selections in the classrooms.

Block schedule issues (5 weeks till DSTP)

Writing Study Group Discussion
Grade 10

Text-based writing is a form of persuasive, expressive, or informative writing – not a separate type.

T's need more practice in developing prompts, etc. for text-based writing.

TM033074

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").